

```
1 ATGAGCCAGC CCAGGCCCCG CTACGTGGTA GACAGAGCCG CATACTCCCT
51 TACCCTCTTC GACGATGAGT TTGAGAAGAA GGACCGGACA TACCCAGTGG
101 GAGAGAAACT TCGCAATGCC TTCAGATGTT CCTCAGCCAA GATCAAAGCT
151 GTGGTGTTTG GGCTGCTGCC TGTGCTCTCC TGGCTCCCA AGTACAAGAT
201 TAAAGACTAC ATCATTCTCTG ACCTGCTCGG TGGACTCAGC GGGGGATCCA
251 TCCAGGTCCC ACAAGGCATG GCATTTGCTC TGCTGGCCAA CCTTCCTGCA
301 TGAGAGGCTG CACGCTTCAG CTTCTTCCCC CTCCTGACCT ACTTCTTCCT
351 GGGGGGTGTT CACCAGATGG TGCCAGGTAC CTTTGCCGTT ATCAGCATCC
401 TGGTGGGTAA CATCTGTCTG CAGCTGGCCC CAGAGTCGAA ATTCCAGGTC
451 TTCAACAATG CCACCAATGA GAGCTATGTG GACACAGCAG CCATGGAGGC
501 TGAGAGGCTG CACAAAAGTA CTACGCTAGC CTGCCTCACC GCCATCATCC
551 AGATGGGTCT GGGCTTCATG CAGTTTGGCT TTGTGGCCAT CTACCTCTCC
601 GAGTCCTTCA TCCGGGGCTT CATGACGGCC GCCGGCCTGC AGATCCTGAT
651 TTCGGTGCTC AAGTACATCT TCGGACTGAC CATCCCCTCC TACACAGGCC
701 CAGGGTCCAT CGTCTTTACC TTCATTGACA TTTTGCAAAA CCTCCCCAC
751 ACCAACATCG CCTCGCTCAT CTTGCTCTC ATCAGCGGTG CCTTCCTGGT
801 GCTGGTGAAG GAGCTCAATG CTCGCTACAT GCACAAGATT CGCTTCCCCA
851 TCCCTACAGA GATGATTGTG GTGGTGGTGG CAACAGCTAT CTCCGGGGGC
901 TGTAAGATGC CCAAAAAGTA TCACATGCAG ATCGTGGGAG AAATCCAACG
951 CGGGTTCCCC ACCCCGGTGT CGCCTGTGGT CTCACAGTGG AAGGACATGA
1001 TAGGCACAGC CTTCTCCCTA GCCATCGTGA GCTACGTCAT CAACCTGGCT
1051 ATGGGCCGGA CCCTGGCCAA CAAGCACGGC TACGACGTGG ATTGCAACCA
1101 GGAGATGATC CGTCTCGGCT GCAGCAACTT CTTTGGCTCC TTCTTTAAAA
1151 TTCATGTCAT TTGCTGTGCG CTTTCTGTCA CTCTGGCTGT GGATGGAGCT
1201 GGAGGAAAAA CCCAGGTGGC CAGCCTGTGT GTGTCTCTGG TGGTGATGAT
1251 CACCATGCTG TCTATCTGTA TCCTCTCCCT AAGTCTGTGC
1301 TAGGAGCCCT GATCGCTGTC AATCTCAAGA ACTCCCTCAA GCAACTCACC
1351 GACCCCTACT ACCTGTGGAG GAAGAGCAAG CTGGACTGTT GCATCTGGGT
1401 AGTGAGCTTC CTCTCCTCCT TCTTCCTCAG CCTGCCCTAT GGTGTGGCAG
1451 TGGGTGTCGC CTTCTCCGTC CTGGTCGTGG TCTTCAGAC TCAGTTTCGA
1501 AATGGCTATG CACTGGCCCA GGTGATGGAC ACTGACATTT ATGTGAATCC
1551 CAAGACCTAT AATAGGGCCC AGGATATCCA GGGGATTAAT ATCATCACGT
1601 ACTGCTCCCC TCTCTACTTT GCCAACTCAG AGATCTTCAG GCAAAAGGTC
1651 ATCGCCAAGA CTGTCTCCCT GCAGGAGCTG CAGCAGGACT TTGAGAATGC
1701 GCCCCCACC GACCCCAACA ACAACCAGAC CCCGGCTAAC GGCACCAGCG
1751 TGTCCTATAT CACCTTCAGC CCTGACAGCT CCTCACCTGC CCAGAGTGAG
1801 CCACCAGCCT CCGCTGAGGC CCCCAGGAG CCCAGTGACA TGCTGGCCAG
1851 CGTCCCAACC TTCGTACCT TCCACACCCT CATCCTGGAC ATGAGTGGAG
1901 TCAGCTTCGT GGACTTGATG GGCATCAAGG CCCTGGCCAA GCTGAGCTCC
1951 ACCTATGGGA AGATCGGCGT GAAGGTCTTC TTGGTGAACA TCCATGCCCA
2001 GGTGTACAAT GACATTAGCC ATGGAGGCGT CTTTGAGGAT GGGAGTCTAG
2051 AATGCAAGCA CGTCTTTCCC AGCATACATG ACGCAGTCCT CTTTGCCAG
2101 GCAAATGCTA GAGACGTGAC CCCAGGACAC AACTTCCAAG GGGCTCCAGG
2151 GGATGCTGAG CTCTCCTTGT ACGACTCAGA GGAGGACATT CGCAGCTACT
2201 GGGACTTAGA GCAGGAGATG TTCGGGAGCA TGTTTCACGC AGAGACCCTG
2251 ACCGCCCTGT GA (SEQ ID NO:1)
```

FEATURES:

Start Codon: 1

Stop Codon: 2260

FIGURE 1A

Docket No.: CL000861CON
Serial No.: TO BE ASSIGNED
Inventors: Karl GUEGLER et al.
Title: ISOLATED HUMAN TRANSPORTER...

HOMOLOGOUS PROTEINS:

Top BLAST Hits:

	Score	E
gb AAF81911.1 AF279265_1 (AF279265) putative anion transpor...	476	e-133
gb AAF71715.1 AF230376_1 (AF230376) prestin [Meriones ungui...	471	e-131
ref NP_000432.1 pendrin [Homo sapiens] >gi 11421915 ref XP...	451	e-125
ref NP_035997.1 Pendred syndrome homolog (human); Pendred'	448	e-124
ref NP_062087.1 Pendred syndrome homolog (human) [Rattus n...	447	e-124
ref NP_067328.1 down-regulated in adenoma [Mus musculus] >...	434	e-120
ref NP_000102.1 down-regulated in adenoma protein [Homo sa...	418	e-115
sp O70531 DTD_RAT SULFATE TRANSPORTER (DIASTROPHIC DYSPLASI...	365	1e-99
ref NP_000103.1 sulfate anion transporter 1; Diastrophic d...	362	1e-98
ref NP_031911.1 diastrophic dysplasia [Mus musculus] >gi 2...	357	4e-97

BLAST to dbEST:

	Score	E
gi 8630793 /dataset=dbest /taxon=960...	523	e-146

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

Expression information from BLAST dbEST hits:

gi|8630793 Human head-neck

Expression information from PCR-based tissue screening panels:

Human fetal lung

FIGURE 1B

```
1 MSQPRPRYVV DRAAYSLTLF DDEFEKKDRT YPVGEKLRNA FRCSSAKIKA
51 VVFGLLPVLS WLPKYIKDY IIPDLLGGLS GGSIQVPQGM AFALLANLPA
101 VNGLYSSFFP LLTYFFLGGV HQMVPGTFAV ISILVGNICL QLAPESKFQV
151 FNNATNESYV DTAAMEAERL HVSATLACLT AIIQMGLGFM QFGFVAIYLS
201 ESFIRGFMTA AGLQILISVL KYIFGLTIPS YTGPGSIVFT FIDICKNLPH
251 TNIASLIFAL ISGAFLVLVK ELNARYMHKI RFPIPTMIV VVVATAISGG
301 CKMPKKYHMQ IVGEIQRGFP TPVSPVVSQW KDMIGTAFSL AIVSYVINLA
351 MGRTLANKHG YDVDSNQEMI ALGCSNFFGS FFKIHVICCA LSVTLAVDGA
401 GGKSQVASLC VSLVVMITML VLGIIYLYPLP KSVLGALIAV NLKNSLKQLT
451 DPYYLWRKSK LDCCIWVSF LSSFFLSLPY GVAVGVAFSV LVVVFQTQFR
501 NGYALAQVMD TDIYVNPITY NRAQDIQGIK IITYCSPLYF ANSEIFRQKV
551 IAKTVSLQEL QQDFENAPPT DPNNNQTPAN GTSVSYITFS PDSSSPAQSE
601 PPASAEAPGE PSDMLASVPP FVTFHTLILD MSGVSFVDLM GIKALAKLSS
651 TYGKIGVKVF LVNIHAQVYN DISHGGVFED GSLECKHVFP SIHDAVLFAQ
701 ANARDVTPGH NFVGAPGDAE LSLYDSEEDI RSYWDLEQEM FGSMFHAETL
751 TAL (SEQ ID NO:2)
```

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

Number of matches: 3

```
1 153-156 NATN
2 156-159 NESY
3 580-583 NGTS
```

[2] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 2

```
1 45-47 SAK
2 445-447 SLK
```

[3] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 11

```
1 18-21 TLFD
2 158-161 SYVD
3 240-243 TFID
4 365-368 SNQE
5 459-462 SKLD
6 556-559 SLQE
7 635-638 SFVD
8 691-694 SIHD
9 722-725 SLYD
10 726-729 SEED
11 732-735 SYWD
```

[4] PDOC00007 PS00007 TYR_PHOSPHO_SITE
Tyrosine kinase phosphorylation site

Number of matches: 2

```
1 7-15 RYVVDRAAY
2 447-454 KQLTDPYY
```

FIGURE 2A

[5] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 10

1	77-82	GGLSGG
2	78-83	GLSGGS
3	89-94	GMAFAL
4	103-108	GLYSSF
5	335-340	GTAFLS
6	435-440	GALIAV
7	481-486	GVAVG
8	485-490	GVAFSV
9	581-586	GTSVSY
10	681-686	GSLECK

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	51	71	0.893	Putative
2	82	102	1.020	Certain
3	107	127	1.729	Certain
4	130	150	1.497	Certain
5	186	206	1.723	Certain
6	228	248	1.517	Certain
7	256	276	1.898	Certain
8	288	308	1.252	Certain
9	338	358	1.568	Certain
10	383	403	1.304	Certain
11	412	432	2.345	Certain
12	469	489	1.997	Certain
13	619	639	1.146	Certain

FIGURE 2B

BLAST Alignment to Top Hit:

>gb|AAF81911.1|AF279265_1 (AF279265) putative anion transporter 1 [Homo sapiens]

Length = 738

Score = 476 bits (1224), Expect = e-133

Identities = 263/724 (36%), Positives = 428/724 (58%), Gaps = 36/724 (4%)

Frame = +3

```

Query: 54  LFDDEFEEKDR--TYPVGEKLRNAFRCSAKIKAVVFGLLPVLSWLPKYKIKDYIIPDLL 227
          L + EE R + P + R +CS A+ A++ LPVL WLP+Y ++D+++ DLL
Sbjct: 15  LNQEHEELGRWGSAPRTHQWRTWLQCSRARAYALLLQHLPVLVWLP RYPVRDWLLGDLL 74

Query: 228  GGLSGGSIQVPQGMFAFALLANLPAVNGLYSSFFPLLTYYFFLGGVHQMPVGTFAVISILVG 407
          GLS +Q+PQG+A+ALLA LP V GLYSSF+P+ YF G + GTFAV+S++VG
Sbjct: 75  SGLSVAIMQLPQGLAYALLAGLPPVFGLYSSFYFVFIYFLFGTSRHISVGTFAVMSVMVG 134

Query: 408  NICLQLAPESKFQVFNNATNESYVDTAAMEAERLHVSATLACLTAIQMGLGFMQFGFVA 587
          ++ LAP+ A N+S ++ A +A R+ V++TL+ L + Q+GLG + FGFV
Sbjct: 135  SVTESLAPQ-----ALNDSMINETARDAARVQVASTLSVLVGLFQVGLGLIHFGFVV 186

Query: 588  IYLSESFIRGFMTAAGLQILISVLKYIFGLTIPSYTGPISIVFTFIDICKNLPHTNIASL 767
          YLSE +RG+ TAA +Q+ +S LKY+FGL + S++GP S+++T +++C LP + + ++
Sbjct: 187  TYLSEPLVRGYTTAAAVQVFVSQ LKYVFGHLHSSHSGPLSLIYTVLEV CWKLPQSKVGTV 246

Query: 768  IFALISGAFLVLVKELNARYMHKIRFPIPTMIVVVVATAISGGCKMPKKYHMQIVGEIQ 947
          + A ++G LV+VK LN + ++ PIP E++ ++ AT IS G + ++ + +VG I
Sbjct: 247  VTAAVAGVVLVVV KLLNDKLQQQLPMPPIGELLTLIGATGISYGMGLKHRFEVDVVGNIIP 306

Query: 948  RGFTPVPSPVVLQWKDMIGTAFSLAIVSYVINLAMGRTLANKHGYDVDSNQEMIALGCSN 1127
          G PV+P + ++G+AF++A+V + I +++G+ A +HGY VDSNQE++ALG SN
Sbjct: 307  AGLVPPVAPNTQLFSKLVGSAFTIAVVGFAIAISLGKIFALRHGYRVDSNQELVALGLSN 366

Query: 1128 FFGSFFKIHVICCALSVT LAVD GAGGKSQVASLCVSLVVMITMLVLGIYLYPLPKSVLGA 1307
          G F+ + C++S +L + GG SQVA SL +++ ++ LG + LPK+VL A
Sbjct: 367  LIGGIFQCFPVSCSMRSLVQESTGGNSQVAGAISSLFILLIIVKLGE LFDLPKAVLAA 426

Query: 1308 LIAVNLKNSLKQLTDPYYLWRKSKLDCCIWVVSFLSSFFLSLPYGVAVGVAFSVLVVVFQ 1487
          +I VNLK L+QL+D LW+ ++ D IW+V+F ++ L+L G+ V V FS+L+VV +
Sbjct: 427  IIVNLKGMRLQLSDMRSLWKANRADLLIWLVTFTATILLNLDLGLVVAVIFSLLL VVR 486

Query: 1488 TQFRNGYALAQMVDTDIYVNP KTYNRAQDIQGIKIITYCSPLYFANSEIF----- 1637
          TQ + L QV DTDIY + Y+ A++++G+K+ + +YFAN+E +
Sbjct: 487  TQMPHYSVLGQVPDTDIYRDVAEYSEAKEVRGVKVRSSATVYFANA EFYSDALKQRCGV 546

Query: 1638 -----RQKVIK--TVSLQELQDDFE-NAPPTDPNNQTPAN-GTSVSYI----- 1760
          ++K++ K + L++LQ++ + P N TS+ +
Sbjct: 547  DVDFLISQKKLLKKQEQLKLKQLQKEEKLKQKQASPKGASVSIN VNTSLED MRSNNVED 606

Query: 1761 -----TFSPDSSSPAQSEPPASAEAPGEPDMLASVPPFVT FHTLILDMSGVSFVDLMGI 1925
          S D A + ++AP + S + A P FH+LILD+ +SFVD + +
Sbjct: 607  CKMMQVSSGDKMEDATANGQEDSKAP-DGSTLKALGLPQPDFHSLILDGALS FVDTVCL 665

Query: 1926 KALAKLSSTYGKIGVKVFLVNIHAQVYNDISHGGVFEDGSLECKHVFPSIHDAVLFAQAN 2105
          K+L + + +I V+V++ H+ V + + G F D S+ KH+F S+HDAV FA +
Sbjct: 666  KSLKNIFHDFREIEVEVYMAACHSPVVSQLEAGHFF-DASITKKHLFASVHDAVTFALQH 724

Query: 2106 ARDV 2117
          R V
Sbjct: 725 PRPV 728 (SEQ ID NO :4)

```

Hmmer search results (Pfam):

Model	Description	Score	E-value	N
PF00916	Sulfate transporter family	254.5	1.5e-72	1
PF00189	Ribosomal protein S3, C-terminal domain.	3.3	8	1

FIGURE 2C

Docket No.: CL000861CON
Serial No.: TO BE ASSIGNED
Inventors: Karl GUEGLER et al.
Title: ISOLATED HUMAN TRANSPORTER...

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00916	1/1	187	497 ..	1	328 [.]	254.5	1.5e-72
PF00189	1/1	651	661 ..	79	89 .]	3.3	8

FIGURE 2D

```
1 CTGGGTTCTT ATGTGGGGAG GTCATGCTCC CCACTCATTG AGCCCCCCA
51 GGCAAACCAC CTGGACAGCC AGACCCATGC AGACTCTGGA GCAGGTGGAG
101 AGGAAAGAGT AGACCACCCC GCCTCACGGG CGGTGAAGGG CCGGCAGCCT
151 CTGAATAGTC TCTGCTAGGA GGTAGAAAGC ACCCTCCCAT CTTAATCATA
201 GTAATCATCG CCACTACCAT TTAATGGGTG CCTATAAAAG GCCAGCCTCT
251 TCATACACAT GATCTCACTG AATCCTCATA GCATCTGCCT GCGACTGTTA
301 CACACTAAGG GCCAGGCGAG AGAAACTGAA TCTTTGAACC CAGGTCACTT
351 GGCTCTCAAA CTTGTGCTGT TTTCCCTAAG CCACCCGGTC TCTCATTTCT
401 CCCACTGAAA TGTCTCACAT GCCATTGCCC TTAATCATTT CTGCCCATGT
451 CTCCTCAAAA ACACCATTTA TCAATTCGCT CAACAAGTAT GTGTTGAGTA
501 CACACTAAGG GCCAGGCGAG GGGCTGGGCA CAGGCGCTGG GGGTAGGTTT
551 ATTCTCCAC CTTGCTTCTT GCTGGGTATC ACCTGTGGGG TCTTGCCGGG
601 CATCCACCCC TCACCTGTAG TTCAAGTGGA CTTGGGATC CCAAGACCAA
651 ATGAATGGAA TGCACAGGCC CAGCCTTCAC CAATTTGAGC ACAATCTTAT
701 TCATAATGAA AACTCACATT TGCATCACAC TTTACATTTT ACACAACCCC
751 TTCTTATCCA TTAATCATTT TGATCTTCAC AACAACCTTG TGAGATATGT
801 CTGTTACTCC CACTTTAGTG ATACAGAAATC TGAGGTTTGA AAAGTAATGC
851 TGACCATTTT GCCTCATTA TAAAGCAGG ATTAACCCAG GCTCCTGGAC
901 CCTTCCAAAG AGGAGCATTA GCAACCTGCT CCCCTCTGAC AACCTCCCTT
951 GTCACCCAGG CTCTCCTCTG GGAAGTTGGG GGCATCTCTA GCCCCCAAGT
1001 AGTTACTCAT TTTCAACCCC ATCTCAAATC TTTTGCCAAA CTGGCCACAG
1051 CCACCCACCA CTCCTCACCT CCCAGATACA AATCCTCACT CTAAGCCTTC
1101 CCCATCTCTT TCTTCTCTGT CTTTCTTTCT CTGTGGTCTT CTGAGCAACT
1151 TCTCCAGCTT CTGGGAGGTA GAGGGGAGGT GGGAGACCCA GTAATTGGAA
1201 GAGGGAGGGG GAAAGGTTCC TACAGGGAAC TCCTCCGGGC CTCAGGGGCC
1251 CTGGCACTCA GCTCTGCCCA TCTCAGCTCC TGGAACGTCA GCCAGGTTGC
1301 GCAAAAAGTG AGGAGGAGAG GAGCGGAGT ACACAAGGGT GGGGGAAAGA
1351 TTAGGCACAG GAAGCCGTGG GAGAGAGAGC CGGCAGGTGG ACCATCCTGG
1401 TTTCCCAACA CACACCATTT TCCCCCTGGG AAACCTGTTG GTGAAGTTCT
1451 AGATGTCTTA TCCAAGAAGG GTCCTCTTGA GGTCATCTCA GCTATCCCCC
1501 TGCCTCTAGG CAAGCTGTTT TCTGTTTCTT CCAAGCTGAC TGGCTGAATG
1551 GTAGGAGCCT TTCTGCCAGG GAAACTAAGG TCTGGGAAGG GAGTATGGCT
1601 TGTGGGGACA CCAGGGGTCA GGGGAGGGGA GGGTCCACCT GCTGAATCAA
1651 GTGGGGCCTC CTGCCCTCGT GATTCCCTTT TGCCTGGTGC TCAGTGGGGG
1701 TGATGGTGAG AAGCAGGTG TGGAGTGCCA GCCACGTGCT GAGCGCCAAG
1751 CAAAACAGCC AGGGTGAGTC TATGCATCAT CAGTGCCTGG GAAGGAAGGC
1801 CACTGCCAGC AGGGAGTCTG ACGGAAAAAC TTGACAGAGG GAAGGGAGGC
1851 ACCTTGCTTT ATCGGGGCGG GGAAGGCCAG AATAAACTC TGCTACTGCA
1901 AGGACAGAG AGAGAAGGCC TGGGCTGGCA CTAGGGAGGG ATGTTCCCTC
1951 ACCCTCCCTT CTTCTGCTTC TCCCAAAGCT TGTAATGCC CCAGATATGA
2001 GCCAGCCAGG GCCCCGCTAC GTGGTAGACA GAGCCGATA CTCCCTTACC
2051 CTCTTCGACG ATGAGTTTGA GAAGAAGGAC CGGACATACC CAGTGGGAGA
2101 GAAAGTTGCG AATGCCCTCA GGTAAGTGGT CCAGAGCCCA GACTTCTGCC
2151 TCCTCTGCTC CTAACCAAAA TCCTTTCTGC ACCAGGACAC GGCTTCTGCA
2201 CTGGTATCCC TAAGATGGGG TTAAGGGAAG CCCTGGGGAA GTGAGGTTCT
2251 GAATGATGAA TTTAAGATCC TACAACCTCA TCTGTACTGA GACCCCAAGG
2301 GAGGATGGGG AGCAGGAGCA AGAACCATCC AGAAGGGTTA TATGGCATTC
2351 CCAAACCCCT GCATGGCATC TCCCATATTC TCAATTCACC CGGGTCTCTC
2401 TGGGTTTGTG AAGGCATGGT AGATGAGCAT CTACGTTATG GAGGGGTGGG
2451 GAGCATCAGA GCCCTTACTC CATGCCCTGT TCCCTCCTTA CAAAAAATAC
2501 CTGAAGTTAC CATCACCCCA GGTTCTTTGT CTTTCCCTC CCGGATGTTT
2551 CTTCTCCAC TTGGTCCAGA GAATGCCAAA AGGAGGCCCT AAATTTCTGA
2601 ACTTCTCTGA GGGGACCTAC CAGGGTGTAG TCCTACCAGC GCCCAGGGTC
2651 TTTCACTCT CATCTCCCTG GAAATGCGAT GGTGGGTATG AAACCTTGTC
2701 CCTAAGTAGG CGCTACACAA GGTGATCCAT ACCCACACC CAGGAGGCTG
2751 GGGCTGCGGG TGTCACCTC CCCATTCCCA GACTCCTGGC AGACCTCCTC
2801 TGGCCAGCT ATAGGCCAAC TCACTCTCCC TCACTCCCTT GGGGAAACGG
2851 CTGATTCAGT TACCTGGATT GAGGTCACTG GCAATGGCTG AAGTGGAGAC
2901 GCAGGTGGAA CTGGTTGAG CCGGGGGAAT CACCCACTTG AGTTTGTACT
2951 AAAAGCCCCA GCCCAGCCCT GTTTCTCTTG GGAGGCTCCA TTTCTGCCCA
3001 GTTACAGTCT GTCCTCACAG CTGTGCTCCT CAGACAGGTG GTCTCTGCCA
3051 GTCTTTGTGC CCAAGACTTT AGGGCACAAA GTCTGAGGAT GAGAAGATCT
3101 GCTATTGTC TAAAGATTA GGATAATGAA AGCTGTAAAG GGATATAGCA
3151 AACTAACAA TCTATGATA CTGGCATGAG AGCCTTGAAC AGTGCCTGGC
3201 ATAGAGAAGG TGCACCAATA AATATTTGTT TCATGAATGA ATGAATGAAT
3251 GAATGTCTAG AAAGCTAATC CCTCTCAGCC TCTGTTTCCA GTTCTTCTTT
3301 CAAGCTTCAG ATTGCTTTGC CCAACATACA GCAGACTTGC AAGTAAGGTT
3351 GGGCATGGAC TAGCCCTCAA ATGAGTTGTT TTTCTTTCCC TAGCCAGCTC
```

FIGURE 3A

3401	TCTATTCATA	AGTCCGGCTT	TCTCTGCCAC	AAACAGACCT	GATGGAGCCC
3451	CTGCAGGGCT	GGTTCTCTCT	TCAAGCAAGG	CTTTAGAGTT	GCATTAAGCA
3501	ATTTATCCCC	CGTCCACCTC	CCCTTCCAGC	ATCCCAGGGA	TGGCAGAGGC
3551	ACCCATGAGC	CCCAGAAGGG	ACAGGGGGTA	AGATATTGAT	GATGATGCTT
3601	TTTCTTGAG	TGTTAGTTGG	AAGAGAAAAT	CTGCCCAGAC	TTTCCAAGGT
3651	ACAAAGCATT	GTCTTTGTTG	GTTTCAGTCT	TGGGTGACAT	CCAGGGGACC
3701	GAGTGTCAGG	GAAACTATTG	TTGAGCAAGA	GCAAAGAGCA	GGAATTGGTG
3751	CTGGGCAGGA	AAGGAAGCCT	CATCAGAGCA	GGCCAGTGAG	TCACCAAATG
3801	GGCCCTAAGT	ATTTGAGTTC	CCTCAACTGG	GAGAAGGAAA	GCAAATGCCC
3851	CTCACCCACT	TCCAGTCATC	AATCCACCGG	CTGTACCCTT	TGAGTTTGTG
3901	AGCCCTTGTT	CCTACCGCTC	CTGAGTTTCT	ATGAAAGGAT	CTTGAGGTGT
3951	TCAACAAACA	GGGAAGGGAT	CAACTCTCCC	CACCCTGCGT	TGACCAATGA
4001	ATTCTTCCCT	CCTCTGCTGC	CCAGTGAATT	AACAGGAGAA	AGAACTCCGG
4051	TATTGGAGTT	ACCACACATA	AAGGATAGTG	AGTCAGCAGA	GTGCACCCTG
4101	CAGGAACAA	AGAGCCTTCC	TTTTCAAGGA	AGTTCTAAGA	AAAATGGCAG
4151	CAGGCAGGCC	CCACTCGGGT	GTATTCATC	ATTCATTTAT	TCAACAAATA
4201	TTTACTAAGT	GCCCCGTGTC	AAGGCTCGAG	GTGTACAAAG	ATGAACAGGA
4251	GAGCTAGACT	TCTTGCCATG	CGTGGTGGGG	TTTGCTGCCT	AGTGGGAGAG
4301	ACAGACAAAA	AGCAAGGAAT	GCACACACAG	GATGCACACA	CAGCGGCAGG
4351	AACCAAGGTG	CAGTTACCCA	GGCCTGGGAT	CAGACAGACA	GGACTCAGAG
4401	GAGACTTTCC	CAGAGAAAAG	CCATCTGAGC	CAAGGGATGG	ATCTGATACC
4451	TCCGAAGGCT	GAGCCACCAT	AACACTCATA	CCTTTAAGCC	AAGTCTTATA
4501	AACTCCCGAG	GTAAGCAGCT	GGCAGTCAGA	AGACCTCCAG	CTAATGCCCA
4551	GGACAAGTTG	ATGAGCTCTC	AAGAAAAAGT	TCCTGCCTTT	TCTTCTCAAT
4601	ATCCCTGGCA	CACAGTTCAG	TGAATTTTGA	ATGAACCAAT	GAATGAAATG
4651	AGCAGGATAT	GATAATCCCT	CTCCAACACG	GAATGTCCAA	GCCATGCAGA
4701	GCCGACTGGA	AATTTTCCCC	GTTCCCTTCC	AGATGTTTCT	CAGCCAAGAT
4751	CAAAGCTGTG	GTGTTTGGGC	TGCTGCCTGT	GCTCTCCTGG	CTCCCCAAGT
4801	ACAAGATTAA	AGACTACATC	ATTCTTGACC	TGCTCGGTGG	ACTCAGCGGG
4851	GGATCCATCC	AGGTCCACA	AGGTGAAGGG	GCTCCTTCAG	CCAGGCTTGG
4901	ATTGCCATCT	CCCTCACCAT	TCCTCTCCTC	ATCCCCACTC	CATCCCTCTG
4951	TGATCCCCAT	AAGCTAGTCA	TGCTGCTGAG	CTTCAGTCTC	GTTGTCCTCT
5001	GCAGGCATGG	CATTTGCTCT	GCTGGCCAAC	CTTCCTGCAG	TCAATGGCCT
5051	CTACTCTCTC	TTCTTCCCCC	TCCTGACCTA	CTTCTTCTCT	GGGGGTGTTT
5101	ACCAGATGGT	GCCAGGTAAG	GCCTCTCCCC	TCTGGGCAGG	CAGGATGACC
5151	CAGACCACAA	GGATGGGAGG	TGTGGCAAAG	GGGCCTCGGG	AGATTTTCCA
5201	TCTGCATTCT	CCTGGAGTTG	TTCTTGGTCA	GTCTAGGGGG	AATGGTCACT
5251	GTGAATGTCA	TTTCCAGTCA	CTCGGTGACC	TTGGAGAAAC	CACTGAGCCT
5301	CTTTGAGTTC	AGTTAGCATT	ACCTGTTCCA	TCTTCTCTCT	AGGAATGAGA
5351	GGAAGACTTA	GCAGAACAAG	ATATACCATTA	TGCTATAACA	TGCTTAAACA
5401	GATGTGAGAA	ATCACCATCT	AACTCCCTGG	TTGGTCCCAG	CCGGCCACTA
5451	CAGGGACATT	TGGACTTCTC	TGGTGCTAAG	TGAGATGGAG	GAAAGCCTGG
5501	TCACAAGGGC	TGGTTTCTGG	TTCAGGCTCT	GCTTATATTT	CTTATTTCTG
5551	AGTTCATTTT	CTCACGTGTC	CTGTATGACA	ATATTGACCA	TTGGGGTAAA
5601	AGCACCTTGA	AAAGCATAGA	TCATGGTTAG	AGTGAGTGGT	TGTTATTATT
5651	GTGTTGGAGA	AGAGCCTTGG	AGGTGCAGGG	ATCCATCCCC	CTGGGGTCGG
5701	GAAGCATTCC	TGGGCCCCCT	TCTGGTTTCC	ATCGGTGTGG	TTCAAACCTC
5751	TGATTTTTGC	TGGCTGGGTG	GGGCACCACA	GGTACCTTTG	CCGTTATCAG
5801	CATCCTGGTG	GGTAACATCT	GTCTGCAGCT	GGCCCCAGAG	TCGAAATTCC
5851	AGGTCTTCAA	CAATGCCACC	AATGAGAGCT	ATGTGGACAC	AGCAGCCATG
5901	GAGGCTGAGA	GGCTGCACGT	GTCAGCTACG	CTAGCCTGCC	TCACTGCCAT
5951	CATCCAGGTG	AGGGGGCAGC	CCCCAACCTT	GCTAGAAGGG	CATCAGACCA
6001	CCCTGCCCC	CCCTCAAAGC	CTTAGCTTTG	ATGCTAAATC	TGATTTAGGG
6051	GGCTGGGTGT	GGAGGCTCAT	GCCTGTAATC	CCAGCACTTT	GGGAGGCTGA
6101	GGAGGGTGGA	TCACTTGAGG	TCAGGAGTTT	GAGACCACCT	TGACCAACGT
6151	GATGAAAACCC	CATCTCTACC	AAAAATACAA	AAATAATCCA	GGCTTGGTAG
6201	TATGCGCCTG	TAGTCCCACC	TACTCAGGAG	GCTGAGGCAG	GAGAATCACT
6251	TGAATCCGGG	AGGCAGAGGT	TGCAGTGAGC	TGAGATCGCG	CCACTGCACT
6301	CCAGCCTGGG	TGACAGAGCG	AGACTCCGTC	TCAAAAAAAA	AAAAA
6351	AAAAA	CCCAAGTTAG	GGCTCACCTC	CTCCCTCCTC	CCCATCCCAG
6401	GGCTAAAGTG	AACCTTGAAA	ATTAACAGTA	TCTCCTCATC	TGCATGTAGC
6451	AGGACCATAC	AAAAAAACAA	CAGCTGTACC	TGGTTAAACT	GTCTGAGCT
6501	TTAAACCTGT	AAAAGACTCA	CAGCCTCTCT	CCATTATCCC	GTGGAGAAAC
6551	CCAACCTCT	GCCAGCATAG	TCTTGACAGC	TGCTAATTTT	CTCTAACATC
6601	CCTCACTCCG	CTCCAGCCTC	CTCTGCTCCA	AGCCACAGCA	GCAGTTGCAC
6651	AACATAAATT	GAGCTTCTGC	AAATGGTTGC	AAAGGATTCT	GCTAGGTTTT
6701	ATGAAGGGAA	GCACAACATG	ACAGAATGCA	AGAGCAAAAC	ACAGTCCCAG
6751	AGAGCGCCTT	TTCATTCACT	CATTCAATCG	GTTTTGTGCC	AAGAACTAGG

FIGURE 3B

6801	CTAAACCCTG	GGATACAAAG	ATAAGTAAGA	AAGAGGTCCA	ATTCACAAGT
6851	TGCTCACAGC	CCAGCAGAGG	AAGGAGCCAT	GTCAACAGAT	AAATTTGTAT
6901	GCAGTGAGAT	AAGCAGCAAA	GTAGAGCCAT	GTACAAAGAC	TGTAGGGACA
6951	CAGAGCAGAG	TCACGGAGGA	CCTCAAAGAG	GAGGTGACAC	TCCACCTCTC
7001	TTAAAGGATG	AGAACTTAAC	CAGGAACAAG	GTATACAGAG	GATGGTCCAG
7051	GCAGAAAGGA	ACAGTGCCTA	AAAACTACTGA	GGCCTGAGAG	AGTGTGATCT
7101	GCGCAGGCAA	AGTAAGGGGC	TTGGTGTGGC	TGGAGGGTAG	AGGGCCCAAG
7151	AGAGGATGGA	AAAGTAGGCA	GGAGCCAGAC	AATGAGATCT	GGGGTCTGTT
7201	CTCTGACAGC	GACTTTGGGT	CTGATTGGCA	GTTTATAAGG	ATCGTTTGGG
7251	CTACACAATG	ATGAGTGGGA	GGTGGATTAG	AATCAAGGCA	GGGGACCTGT
7301	TGGGAGACTC	TGCAGAGGCC	CAGGCAGGAA	TAATGCAGGC	GAAGACCAGG
7351	TAGAGAAAGA	GATGGGGCTG	GACTTGAAAA	GAATGTTTTA	CCAGGAGCTT
7401	GGTGATAGAC	TGGATGTGGG	AGGTAAGGGA	GGATGACTCT	CAAGTTTTTG
7451	GTTGGGCAAC	CAGGTTAATG	ATGGTGTCA	TTACTGAGAG	AGAAAACT
7501	GGGGGAGGAC	TAGACTTATT	TTACAGATAA	GCCAAAGCCA	GAGAGGTGAT
7551	GTGACAGAAA	GGCCCATGCT	CTAAAGGAGC	TGAAGGTCTG	ATGGCAGCCA
7601	TGTAGAGCAC	AGTGAAGGGC	AGGTGAAGGT	CACAGATGGT	CCAATTCCT
7651	CAAGCTACTG	CTACGCTAGG	ACTGCACGGA	GCTCCAGACC	TGCGTGTGTG
7701	TGGGGCGGGT	CGTTGGAAC	GCTGAACCA	ATTGGTCTTC	CGCCACCAAC
7751	CACCCTTTTC	CTCCTCTCAG	ATGGGTCTGG	GCTTCATGCA	GTTTGGCTTT
7801	GTGGCCATCT	ACCTCTCCGA	GTCTTTCATC	CGGGGCTTCA	TGACGGCCGC
7851	CGGCTGTCAG	ATCCTGATT	CGGTGCTCAA	GTACATCTTC	GGACTGACCA
7901	TCCCCTCCTA	CACAGGCCCA	GGGTCCATCG	TCTTTGTGAG	TCTGGGGATG
7951	CACCCCTGCC	ATTGGAGCAA	GGCTCCAGCA	GACACATGAG	GAGGATGTAC
8001	TGTTTTAAGA	TGTCGTGAGC	TCCTCATTGC	AAGGGCTGGC	TTAGCTGTTG
8051	TTCAGAGAGG	ATTCTGAGGG	GGTTTCTGTC	TTGGGAGGGT	CAAAGTCATG
8101	ACTCACAGAG	GTTCTTGTA	GTTAATACCT	GCAGAAAAGA	GCTGTACATT
8151	CTCCGCCAGT	TCCCCATTCT	AGTGCCTCAA	CCCCTCCCTG	CCTGGAAGT
8201	CCTGCCTTAT	GTCTAATCTC	CATCCCTCCT	CCTTCAGCCC	AAACTCTTCT
8251	AAAGAAAAAG	AAAGCATTCC	TTTTCTAGCA	CAAGTTCCTC	ATGTGCCTTT
8301	TGGGAAAGGG	CGGTGGGCGA	CGGGACAGGG	TTCTGATCA	GGGTTTTAAT
8351	TCTGTCTTGG	TGTGCCTCCA	TTAGCTTTGA	TGGCATCCCT	TCCCTGGGTC
8401	AGACACCCAA	AGGTGGGGTA	TTATGGGAAG	AAGGGGTGGG	AGCCTGTGAG
8451	CATGATGCTC	TTTCCCCAG	ACCTTCATTG	ACATTTGCAA	AAACCTCCCC
8501	CACACCAACA	TGCCTCGCT	CATCTTCGCT	CTCATCAGCG	GTGCCTTCCT
8551	GGTGCTGGTG	AAGGAGCTCA	ATGCTCGCTA	CATGCACAAG	ATTCGCTTCC
8601	CCATCCCTAC	AGAGATGATT	GTGGTAAGGA	CCTTGTTTCA	AGCTGGGATG
8651	TTGGGGGGGC	AGGCTGTGAG	ACGAGGAAGC	CCCTACCTTT	CCTCACCCCA
8701	TCCCCTCAAC	TGGCAGCCAG	TGGGACAGGA	AGTCAGTTGT	GAATCCATCC
8751	CATCCCCCGT	ATGTGGCGTT	TCCTCTCTTT	CTACTGCTCT	AATAATTCCC
8801	CCTAAGGAGG	CAGGGGAGTG	GGATTCAAGG	TCCCCAGAGA	AAAGGGAGAC
8851	TTGAGAGAGA	CGCTGCCCT	GGCCCCACCT	TAGGGCCAAT	CCCCATTCTC
8901	CACTCTGGGG	TTTGCAGGTG	GTGGTGGCAA	CAGCTATCTC	CGGGGGCTGT
8951	AAGATGCCCA	AAAAGTATCA	CATGCAGATC	GTGGGAGAAA	TCCAACGCGG
9001	GTGAGTCCAG	GTGGCCCAAG	AGCCTGGCCC	ACCCGCACCT	CATGCCCCAC
9051	TAAGGCTGAT	GCTCGGAGAG	GGAGACAAGA	TGAACCTCTA	GAAAGTGACG
9101	TCGAAACTGT	ATGACACTGA	CCATGTATGA	ATTATTACTA	TTACCGTTTC
9151	CTGAGAAGGG	CCGCACAACC	AGCCAATGTA	GGCTATTTTA	TGAGAAATGA
9201	GTCTTAACTG	CCACACTCCC	CTTATAAATC	TCAATCAACT	GATGCTGTTA
9251	AACAAAGCCT	CTCTGAACAG	CCGCTTGCTG	GCTCTTTGCC	TTGCTCTAAT
9301	GCATTGGTTC	TTTGTCCATG	TAGAAAGGGA	ACTATTAGGT	TCAACCAGAT
9351	TCATGAAGCA	TCCACTCTGT	GCCAGGCACC	ATGCTGGGCC	CTGGGAGGAG
9401	AGGGGTGACG	CTTGTCTGTC	AGGGTTGGAA	CAGGCAAGGG	AGGGAAGACC
9451	ACATAGCACC	AAAGGTCTAG	GGGTCTGTGG	ACTCGTGAGC	ATACAGGGTT
9501	CAGAATCTGG	GAGTTAAACA	ACGAGGCCCT	ACCACATACT	GGCCCCGGGA
9551	CCTTGGGCAA	GTTAGGTTCT	CTCAGCCTCA	GTTTCCTCCT	TTGTAAACA
9601	GGAGTGATGG	TCCCTACCCT	ATGGGGTGGT	GCTGAGGATT	CAGACTGGAT
9651	GGGATAACTT	AGGCAAAGAT	CCCGGCACAC	CATGGGGGCC	TGGCTGGTCC
9701	CTGTGGGCTG	GTGAAGGACT	TGGCTGCCCT	CCCCACTCAC	ACCTTGGGT
9751	TCTGCCTCCT	TCCTGGCTCC	TCGGCAGGTT	CCCCACCCCG	GTGTCGCCTG
9801	TGGTCTCACA	TGGAAAGGAC	ATGATAGGCA	CAGCCTTCTC	CCTAGCCATC
9851	GTGAGCTACG	TCATCAACCT	GGCTATGGGC	CGGACCCTGG	CCAACAAGCA
9901	CGGCTACGAC	GTGGATTGCA	ACCAGGTAGC	TCTGGCCACC	CCCGGCAGGA
9951	CTGGGCAGGA	CAGGTCAACT	CAGGCCTGGC	ATGACATATC	TTGGGTGGGG
10001	AGATCATTGG	GCTGAGGTGA	GGCAGGCTGC	CTCGAGTGTG	GGGGATAGGG
10051	GGTCTCTGTA	CCCTAAGAGG	CTGACCTCCT	CTTGACTGGG	AATGTGTGAC
10101	TTTATAGCCA	CTGGGTCACT	CTCAGGTCTT	AGGCCACAG	TCCAGCTTGC
10151	ATGCCTGACT	GCACTTGCTC	CCCGTGCCCC	CCAGCCCCAC	ACTGGCTTCT

FIGURE 3C

10201	AATCCTGTCC	CCTCCCTGCA	GGAGATGATC	GCTCTCGGCT	GCAGCAACTT
10251	CTTTGGCTCC	TTCTTTAAAA	TTCATGTGAT	TTGCTGTGCG	CTTTCTGTCA
10301	CTCTGGCTGT	GGATGGAGCT	GGAGGAAAAT	CCCAGGTGAG	CCTTGTTCTA
10351	GGGGAGTTGG	GGGGAGGTGG	TAAGAGAAAC	GTTGCCCCAA	AAAAGCCTGG
10401	GCACTGCAAG	CCAGGCCAGC	TCTTCTCCGA	CCCCTTCTTC	CCGTACTTAG
10451	TCTCCTCTCC	ACCAAAGCCA	TGGATTGGAA	ATAAATCAAG	AGCAAAAAAT
10501	TCACACCTTC	CCTCTATCCC	CAACTCTTTC	TCGGAATAGG	TGGCCAGCCT
10551	GTGTGTGTCT	CTGGTGGTGA	TGATCACCAT	GCTGGTCTGT	GGGATCTATC
10601	TGTATCCTCT	CCCTAAGGTA	AGAGCCCAGC	CATCGAGCAG	AAGTCAACGA
10651	AAGACTCCAA	TAAGAACAAAT	CCCTGAGAGT	TGTGTGGCAC	TTTACGGACC
10701	ACAAAGTGCC	ACTGTTGTCA	TACTTAGTCT	CAACCACAAA	CTGTGAGGTA
10751	GACAATGCAG	GTTTTATCCT	CCCCATTTTA	CAGGTGAAGG	AAACTGAGTC
10801	TGAGAGTCTA	AGTAACCTTG	TCCATAGTGA	GGCAGCTTAC	AGCGCAGGGC
10851	TGGTCCCAAA	CTCCAGCCTT	CTGGCCTCAG	AGTCTAATCC	CTAGGCAACA
10901	TTTGCCAGCT	CCACGAGTA	CCAGGCTCTT	ATATAGCCCA	GCTAGGAGGG
10951	CTCTAGGCAT	GCGTCATTTA	GAGATGAGGG	AAGAGAGATA	GGGAAAGGAT
11001	GGGGCCAGGA	AGGACCCCAT	GGCTCTAACG	CCAGCACTTT	CCAAACCTAA
11051	GGTCAATGTC	AGAGATTTTG	GGGATCAGCC	AGGGGAGGTG	TTCCAGAACT
11101	CCGTCTCTGT	CTCGCCAGGC	CTTGGGGTCG	GGTATGCGCA	GGAGGGCAAA
11151	AAGAAGGGGA	GACCCTGGGG	TCTGGAGCA	ATGTTCTGCT	TCTCTAGTCT
11201	GTGCTAGGAG	CCCTGATCGC	TGTCAATCTC	AAGAACTCCC	TCAAGCAACT
11251	CACCGACCCC	TACTACCTGT	GGAGGAAGAG	CAAGCTGGAC	TGTGTAAGTA
11301	TCGGGCAGCC	TCTGGGTACT	GGCCATGCCC	CTGCCCTCTC	CTCCAACCCC
11351	ACAGCCCTGT	CAGCCCTGTC	CTAACAATGA	ACCCTCTAGT	CTGCTGCTTC
11401	CTAATTAGCA	TGAGATGAGT	GGTTAAAAAGT	CCGAGTTTCG	AAGTGAACA
11451	TCCTATGTTT	AAACCCCTAAC	TCAGCCATCT	GCTGGCTCCA	TGGCCAATAG
11501	CAAGCCCCTT	AACCTTTCCC	AGTCTTGTTG	TCTTAACTGG	GCAAATGGTT
11551	ATTTTATGCT	CTCTGCCTCC	CAGGGTTTTT	TATGAAGAAG	AAGCAAGGTA
11601	ATACAAGTAA	ACATGTTGTC	TACATCGTAT	TTTATACTCA	ATAAAGCTTA
11651	GCTATGACTA	CTTTATGACA	TACAGCTTTA	AAAAACAAAA	GGAAATAGTT
11701	TGTATTTTAA	AAAAAAACCT	AGAACATAAA	GCCAGAGGAC	CAAAATCTTG
11751	AGCAAGTTAC	TAGACTTCCC	TGGGGTTCTA	TTTCCTCATC	TGTAAATGGG
11801	GGTGAGACTC	ATGCAGTCAT	GGTTGCGTCA	AACGCTGGTT	CCGAGGATTA
11851	AATGAGATCC	CAGTGGGAAA	ACACCGCATG	AGCGCAAAAC	TTCTGCAAAC
11901	ATGACTTATT	GTCCTGATTA	GTCACACACT	CCACCGCATC	ATCCGCTGGG
11951	CATAGTAATG	AAGGCCAGTG	TGTTTTGACG	ACACTGCCTT	CTCTCCATTT
12001	AAGCCCCACC	ATAACCTATG	GGAGAGGATT	TACTAAACTT	TCTTAACGGT
12051	GATGAAACCA	AGGCTCAGAA	TGGTTAAGTA	AATTGTCAAA	GGCCACAGAG
12101	GTAGGGAGTG	GTAGAGTCTG	GATTAAAAAT	CCAAGTCCTG	GACTCCAGAC
12151	CTCTAGGCTG	TACTGTCTCA	TAGGGAAGGC	AGTCTCACCC	ACCTAGGGCA
12201	GAGAAGAAAA	TCCTTAAAGC	CAGAGAAGTG	AGTGGCTCAT	CTGTGGTCAC
12251	CCAGAGAGAC	AGTGATGAGG	ACAGGGAGAA	AAATTATACC	TCAGTTCCCA
12301	GCCCCAGGAT	CTGCTTTGAC	CATAACCCAA	CAAGCCCCG	CTATGGTGGT
12351	ATTTCTTAG	GTTTCATATG	CGGCTTTTGT	TTCCATTTGA	TCTTCACAGC
12401	AATTCTCTAC	AGGAATCTGG	GCAGATTTAT	TTCTTTTAGA	GGAATTTCCA
12451	GGTCTTAAAA	TCTATAGGGG	GCAACTATCA	AAACTTCACC	CAATGTTGCC
12501	CCCTACCCAC	ACACAAAACC	AGGCCCCAG	CCGATCAGAA	AGCACTGCTG
12551	AGCTCCTGTC	AGGGCCACG	CAGCTCGCTG	TGAGACAGAG	AGAGGGAACT
12601	CACATTTATT	GATCACCTAC	TGAGCATCCA	TCACTAGGCT	AGGACCGTCA
12651	CATTCCTTAA	CTTTTGAATC	CTTTTCATGAG	GTAGGCATTA	TTATTCTCCT
12701	TTTGTTCAC	ATAGCCATTA	AAGAACAAAA	TTTGGGGCTG	GGTGTGCTGA
12751	CTCACACCTG	TGATCTAGCA	CTTTAGGGGG	CTGAGGCAGG	AGGATCGCTT
12801	GAAGTCAGGA	TTTCAAGGTC	AGCTTGGGCA	GCTTAGCGAG	AGCCGTCTCT
12851	AGAAAAATAT	AAAAGTTAGC	TGGGTGTGGT	GGCACGTGCC	TATAGTCCTA
12901	ACTATTACAG	AAGGTTAGGC	GGGAGCACAA	CTTGGGTTCC	AGGGTTTGAG
12951	GCTCCAGTGA	GCTGATCTTG	CCACTGCACT	ACAGCCTGAG	CAACAGAGCA
13001	AGACCCTGTG	ACTCCAAAAA	CAAAACAAAC	AACACATTTT	GAACCCAAAC
13051	AGATCTGACC	CAAGATGCAT	GCTCTTATAG	ATGCCACCTC	CCTGTGTGCT
13101	GGGGCTTCTA	CTAAAAACAC	AGACAAGATC	AGGCAACCAC	AGTCAATCTA
13151	AGGGAAAGAG	GAAAGTGTA	CCAAAGCACA	AATACATAAA	ATATTGCAAA
13201	AATGCTATTT	AAAGAAAAAA	AAGAGAAGAG	AGGCTCTGAG	GTTGTAATAA
13251	CAGAGAATGG	CCTTGGCTAA	TCCAGGAAGA	CTTCTGAAA	GAGGTTGTTT
13301	TTTCCCAAG	TCTGCTTTTG	ACATCTCTCT	TTTCACAGTG	CATCTGGGTA
13351	GTGAGCTTCC	TCTCCTCCTT	CTTCTCTCAG	CTGCCCTATG	GTGTGGCAGT
13401	GGGTGTGCGC	TTCTCCGTCC	TGGTCGTGGT	CTTCCAGACT	CAGTTGTAAG
13451	TGATAGCTTC	GGCCCTCCTA	GGCCACAGT	CGGTTCCCTG	GGCCAGCCCG
13501	CAAGGGGCTT	CCATGCCACG	GCCTGGCTTA	GTCCACTGTA	CCTTCCACCT
13551	CTGGGCCTGG	CACTGGAGGT	GCTGCCAGGC	CCAAAGAGAG	CCCAACCCAG

FIGURE 3D

13601	CCAGGACTGT	GGGCACAGTC	TGGGCTGTTT	GA CTTC C C C A T	A T C T T G A A A A
13651	CCCCAGAGAA	AGCCAGCATA	CTCTTGCTGG	GGATGGCTGG	GGAGAGGGCA
13701	GTGGCAGAGA	AAGGAGGGCA	AGGGCAGGTG	GTGAGATTCA	ACATCCTTCC
13751	AAAGACATTG	CCAGAACCCC	AAACCAAATG	GGACCCCACT	CCAGGAGAGC
13801	GCCAGGGTGG	AAGACAGAAG	CTGTGTTCTA	CACACTGGGA	GTATTACAGA
13851	GAAAGGGTCT	TGGCCAAGGC	AGGGAGTACG	CTGAATGTTG	GGGGAATCCT
13901	ATCTTCTCTT	CTTGAGAACT	CAGAACAAGG	AAATGATGAC	TTCAGGGCGA
13951	CTCCCACCAC	TTCTCCCACC	ACTTCTCTCC	CCTGCCCTGT	GGTCTGGGAG
14001	CTATGTCAAG	GACCTGCCTG	TCATCCTCAT	AGTTATAGGA	GGCCACAGGC
14051	CACCAGACAT	GTGTCTCCAG	TGCAAAAAGA	CAGACACAGC	AAGTCTGGGG
14101	GTGAGGACAG	GACCCCATCC	TACCTTGGCT	CTGCCCCCGC	CCCAGCAGGG
14151	GCACCCCTTC	AGGCCCATGT	GCCATTAGCA	TTCTCTTATG	TTTTTCTCTT
14201	CCTGCTTCAT	CCAGTCGAAA	TGGCTATGCA	CTGGCCCAAG	TCATGGACAC
14251	TGACATTTAT	GTGAATCCCA	AGACCTATAA	TAGGGTAGGT	AATTCAAGCT
14301	TATGACCTTC	TTCTTTTGCT	CTGCACCACC	CCAAGAAGAG	GTTGCTTTTT
14351	AAAGCCAATA	AAGACATTTT	TGCAACTTGA	GCTCAGTCTC	CCTGTCACAG
14401	GCCCAGGATA	TCCAGGGGAT	TAAAATCATC	ACGTACTGCT	CCCCTCTCTA
14451	CTTTGCCAAC	TCAGAGATCT	TCAGGCAAAA	GGTCATCGCC	AAGGTAAGGC
14501	TCAGTCCCTG	GACCCACAGG	GCTCTGGACA	GAGAGTGGCC	GGAAAATGGA
14551	AGCAGAAGGG	CGGTGGGAGC	TGAGAATAGG	CCACTCCCAT	AGAGGGTGGA
14601	GGTCAAGATT	GCTGTTGGCT	CTCTCCCTGC	AGACAGGCAT	GGACCCCCAG
14651	AAAGTATTAC	TAGCCAAGCA	AAAATACCTC	AAGAAGCAGG	AGAAGCGGAG
14701	AATGAGGCCC	ACACAACAGA	GGAGGTCTCT	ATTCATGAAA	ACCAAGGTGA
14751	ATGAAGGCCA	GAAGCAGCCC	CGTGCCCTGC	TCTCCTGCCC	ATTCTGATAC
14801	TGCCCCCTGT	TACTCATGGT	ACCCTGGGGG	CCCCGCTTCC	CACCCTGACA
14851	GGCAAAGACA	GAAAGTCTCT	GGGAACACTG	CCTGGTGGCC	GCTGGGCATT
14901	TTTCTTCTTT	TTTTCTTTT	TCTTTTATGA	GATGGAATTT	TGCTCTTGTC
14951	ACCCAGGCTT	GAGTGCAATG	GCGTTATCTT	GGCTCACTGC	AACCTCCACC
15001	TCTGGGGTTC	AAGCGATTCT	CCTGCCTTAG	CCTCCCAAGT	CGCTGAGATT
15051	ACAGGTGCCA	CCACACCCAG	CTAATTTTTG	TATTTTTAGT	AGATATTGGG
15101	TTTCACCATG	TTGGCCAGGC	TGGTGTCAAA	CTCCTGACCT	CAGGTGATCC
15151	ACCTACCTTA	GCCTTCCAAA	GTGCTGGGAT	TACAAGCCTG	AGCCACTGCG
15201	CCCAGCCTGG	GCATTTTTCT	TCTTGGATGA	GGTGCTACCA	TCTCCCAGGG
15251	AAGCCACTGA	ACCCCCAAGG	CCCTTCTCCA	TTTTCTGGCT	AAGATAGGAC
15301	ATGGCCCCTT	GATTTTGAAG	CAACCCAGAG	GGGGAACAGC	AGTGAATTTT
15351	CTGGGGAACC	CAGGCAGCCC	AGGGCTAGCA	AGGCTGGGGT	GGCCATGGCA
15401	GTAATCCTTG	TAATCCCAGC	ACTTTAGGAG	GCCGAGATGG	GAGAATCACT
15451	CTCATGAGTT	CAGGAGTTCT	AGACCAGCCT	GCCCAACGTG	GCGAAACGCT
15501	GTCTCTACTA	AAAATACACA	AAAATTAGCC	AGGCGTGGTG	GTGGGCACCT
15551	GTAATCCAG	CTACTCAGGA	GGCTGAGGCA	CGAGAATCAC	TTGAACCCGG
15601	GAGGCAGAGG	TTGCAGTGAG	CCGAGATAGT	GCCACTGCAC	TCCAGCCTAG
15651	GCAACAGAGG	GAGACTCTGT	CTCAAGAAAT	AAAGGAGCTC	AGTGTCCCCG
15701	GAGGGGCTTT	CTCCAGAGA	GAGTGGGCTT	GAGGCTTCAG	TGCTCTCTTT
15751	GGCTGGGTCC	TCTGACTTTG	TCTGGGTTGT	AGGAGACCAA	GTTTGCAGGC
15801	CCTGCCTAAG	AAAGGGCTTT	GGGAGAGGCC	TCTCTGGTGG	AGCTTTCAGG
15851	GTCTGTGTTT	ACCATCACCG	AGGCGAGTTA	TTCCCCTACA	CCTACACCCT
15901	CCATGCCCTT	GCTTCAGTCA	CAGCAAGGTC	TGGCTCAGTC	TGGTGGTCCC
15951	TGACTCTGCC	CAGTGTCCCC	ACCCTTCCAG	ACTGTCTCCC	TGCAGGAGCT
16001	GCAGCAGGAC	TTTGAGAATG	CGCCCCCAC	CGACCCCAAC	AACAACCAGA
16051	CCCCGGCTAA	CGGCACCAGC	GTGTCCTATA	TCACCTTCAG	CCCTGACAGC
16101	TCCTCACCTG	CCCAGAGTGA	GCCACCAGCC	TCCGCTGAGG	CCCCCGGCGA
16151	GCCAGTGAC	ATGCTGGCCA	GCGTCCCACC	CTTCGTCACC	TTCCACACCC
16201	TCATCCTGGA	CATGAGTGGA	GTCAGCTTCG	TGGACTTGAT	GGGCATCAAG
16251	GCCCTGGCCA	AGGTGAGGCC	CTCGGGGACA	GCAAGCACCA	CCCCTCCAC
16301	CCCCTGGCCT	CTGCTCTCCA	CATTCCCTTT	CCTGGGAGCC	CTCATTTTCA
16351	GAAGCTGAGG	GAGGAAGCTC	ACTGGGGAGA	CTAACAGCTC	CTAGGAATCC
16401	CTCCTTTCCC	CAGACGCCAC	CAGGTTGAGA	CATTCTCCAC	AGAGCAGGCC
16451	CAGACGGCCC	ATGACAATGA	GTGGCGGGAC	AAGTCTACCA	GAGTTTCAGG
16501	CCCCTGTGCT	CCCAACACCC	CCAGCAGTGG	CCATCCCAAG	TCCCTCTCAG
16551	CCATCAGGAA	CCCACCCAGG	TTCTCTGAGG	AGGGTCCAGT	TTGGCTCCTG
16601	GTTTCATGATC	TGCTGCCCTT	GTCCCTCATT	CACCAGCCAC	CCTAGGACAG
16651	GAGAAGAAAT	AATACCAAGT	CCCCACACCA	TCAGGCCAAA	CAGAGAGCCC
16701	ACGGGACACC	TTGAATGAAT	GTATCCATCT	GATAACTTTC	CAGCAGCCAC
16751	CGCCAATGGC	GGGAGTCAGC	AAACCTCAGA	GCTGGCTCAG	ATAGAGGCAA
16801	GCCAGGGGAA	CAATGGGCAC	AGAGAGTGTT	CGGACTGCCT	TCACCATCAA
16851	CCAGGCGCAG	GGCAGGCCCC	ATACCCAGCC	TTGGGCCTCA	GCCGGCTTCC
16901	TTAGCTCAGG	TCTGGAGTCC	AGGCCAGCCT	TGGCTGAAGC	TCTAGACTCC
16951	CTGAGCCTCC	ATCCTCCCCT	GCAGCTTCTG	TCTGAAGCCA	CAAAGAAGTC

FIGURE 3E

17001 TGAGAATCTA AGCTACTGAA AGAAAAGATC AGCCGGGCGT GGTGGCTCAC
17051 TCCTGTAATC CCAGCACTTT GGGAGGCCAA GGCAGGTGGA TCACAAGGTG
17101 AGGAGTTCAA GACCAGCTTG GCCAACATGG TGAAACCCCG CCTCTACTAA
17151 AAATACAAAA ATTAGCCAGG TGTGGTGACG GGCCCTGTGA GTCCCAGCTA
17201 CTCGGTAGGC TGAGGCAGAG AATTGCTTGA ACCCAGGAGG CGGAGGTTGC
17251 AGTGAGCCAA GATCGCGCCA CTGCACTCCA GCCTGGGCAA CAGAGTGAAA
17301 CTCCATCTCA AAAGAAAAAA AAAGAAAAATA TCTAGCCCCA CAAGAAGGGG
17351 CCATGGTGAC TTAAAGTGCC CGCCACGTTG GCAAAAGTCC ATTTCCGCTC
17401 CACTTCCCAG AGAAACCGTC AGCCAACT CCAGGGAGAA GTGGTGTGCT
17451 TTGCTGCTAT TTTTGTCTTT GGCTGCTGGG CTCTCAGGGT TGCTTATTTG
17501 TTTGGCTTCC CCTCTGAAGT ACGTTTTGTG AATCACTTTT GAGACCCACT
17551 CAGAACATTC CTTTCCTTTT GCCTCCCTAC CCCAACACA CTTCTAGCTG
17601 AGCTCCACCT ATGGGAAGAT CGGCGTGAAG GTCTTCTTGG TGAACATCCA
17651 TGGTAAGAGA AAGAGGACAT TTAGGGACTG AAAGACTGGC AAGGAGTGTG
17701 GGGTAGGAAC AGGTTGGTGG GGTCTGAATA GTGAGGAGGT TGGAAACGAG
17751 AGCACCCAGC TATCCCCAC AAGCTGCTGC CTGCTCATAA AAGCTTCAGG
17801 TACAAGTCCA AAGAGACTGG TCAGATTGCA TAAACATCCT AGGGGCTTFA
17851 GTGACAGAGT GGGGGTGAGG AGGTCATGGA GTTACAGAAG GACAGCTAGG
17901 ATTCTAATCT ACCCCATAAC TAATTTGCCA CGTATCCTTG GCCGAGTCAC
17951 TTTATCTCTC AAGGGATCTA TTTCTACCTA TGTAAAACGA GAGGGTTGAC
18001 TAGATGGATT TGGGGATCCT CTCCCAATCA GAAACTCTGT GAATCGATAT
18051 AGGCATAGAG CACACGGTAC CCTAATTCCC CAGGGAACAT ATAAATATGC
18101 AGTTTTGTAG GCATACAGCC TCCAAAGGGT GCATATACAC AGCCTCAAGG
18151 ACGTGGCCAC AGGGCAGCAG ACATTTACAT GACTAGCATG TACGCAAAGT
18201 GCAGAGATGT GGGAGCAAGT GCACACAGAC ACACAGGAGA ATGTGAAGGG
18251 GCACATACAC ACACACCCAG CTCCCTGCAC TGGGTGAGC CCCCTCCAGC
18301 AGGGCTGCAG TTCCCAAGCT CCGCATGGCC ACGTTCGGGG AGAGAATCTG
18351 CAGTGGCAAT GACCTGCTAT GATATGTTCT GGAGTTAGAA GCAGTGGATT
18401 CTCCCCAACC TCACTGGACA CCCCCTTAGG AAACCATCTC TAGGATTAAG
18451 AGTAATCCAC ACAAACTTCC AATGCCACAC ATTGGAAGTT GCTGGAAAGG
18501 TCTGGGAAAA CAAGAGGAAG GATGGGTCCT TGGGGGATAG AACTGGCAGC
18551 GGCCTCTTCA AGGATGGCTT AGGCTTTTCC ACTCGAATCA CCACAAAGTA
18601 CTGACTCCCT AAATCAAACT GCTTCCTTCT GCTCTGGGTT GAAACTTCAG
18651 CATCCTCAAG TTCTATGTTG CCTCTGCCGT CCAGAACTGA TATTGCACTG
18701 CCAATGCCAT GGCCCTCAGA TACAGCAAGA GCTGGGACCT CAGGCCTCTC
18751 CCATCCCTGC TCTGGTCTCA CTATCTTCCC CACCCCGAGC TCCAATCCAC
18801 AATGGCTGTT ATCTTTCTGA AGGTGATCTT TTCTCCTTCT AGCCAGGTG
18851 TACAATGACA TTAGCATGG AGGCGTCTTT GAGGATGGGA GTCTAGAATG
18901 CAAGCACGTC TTTCCAGCA TACATGACGC AGTCCTCTTT GCCCAGGCAA
18951 ATGCTAGAGA CGTGACCCCA GGACACAAC TCCAAGGGGT AAGGTTCTTG
19001 CACCTGGGGA ATCCTAGGCT CCAAGGCACT GAAATAGCAG GACCAAGAGG
19051 CATTATTAGA AAGAACACAG GAGAAGGTTT AAGTTCCAAT ATCAAGTCTG
19101 CCATTTCACT TTTCTGAATC TGTTTCCTTA TCTATAGAAT GAGCACCATC
19151 AACTAACATT ACCTACCTCT CTGCATTTTT CTTTTATTTT GTTTTAGGGT
19201 TAAATGATAA TTAACTCTTT TGTGTCACTT GAAAGCACTT TGTGTATTGT
19251 AAAAAATCTT TATCAATATA AGTTTTCTGG TTGCACAAAC ACCCAAAGCA
19301 TAGTAGAGCA GGCCCACTCT GCTGGCATCG TTCCCTGCCT CCTCCTCATC
19351 TCTTTCTAAA GGGGGCTTTC GGGGAAGGAG GGGAGGGGAG TAAGCCTACC
19401 CATTTTAACT TACCGGAGCT TAGAGATTTC AGGCTGGTGA GGGATAAAGA
19451 GATTGGGTCT GAGTTTTGTC TCAGCTTTTT GACATTTAAT TTAAGTCTC
19501 AGTAAGTCAT ACAAATGGGA TACAAATAAC ACCATCTAAA ACTCCAGAAG
19551 ACTGGGGAGT CAGAAAAATC CTACCTCCTT GGGGTCCTG CCCAGATCCC
19601 CAGTCATCTC TAGCCCTCAG GGTCCCCTCC CAGCTCAGT CCTGCCCTTG
19651 GCCTCCCAAG ACTCTTGTG TGCCCCAGCC CTGGGTAAAA ACCTCCCCTG
19701 CCCTCTGTGG GTCATAAGAA AGGCTTTTCT GGCCCTAGAG CAATGATTTG
19751 CTCTTTGCCT TAAGAGACTG ATGAAGGTGA AACCATCTGT TCTAAGTGCT
19801 GAAAGACTGC CCAGGAACAC ACAGGGCGCT GGCTCCTGCC CTCCATGCCT
19851 AGAGGGGAAAC CCTGGGGAAA CAACGGGCTT TCCTGCTTCG TGAAATTTGT
19901 CCGCAGAGCA AAGAGGGAGA TTCTGGAGGA AGCTGCATTA GTTGTTAGTG
19951 CCCTAATCAT GTTCAGCTAC TCTAGTTGGT ATGTATACTT GATTAGTCAT
20001 AGCACTTATA AATAATTTAT ATTTTATATA ATATATACTT ACATATTATA
20051 GACCATTACAG AGATACAAAT CACACACATA AACACACACC TTTTCAACAG
20101 CATTGTGAGG GACAAAGCAG GCAAAGTGAG GCTGGTTATC AGACTTTAAC
20151 AGATTAGAAA ATATATTCCC AGGAGGACAG GAATTCCTCA AGGTGAGGCA
20201 GCTAGCCAAT AGTTTTCTA AGCTGAGTAA AACCTTCCCT GCCTCTAACG
20251 GCCCACAAAG GAGGGAAGAC CGCGATACAC ACCTGTCTGG TATAAGGGGG
20301 AAGACCACAG CCGTGCTGTT TTTGTGAGGC AGGTAAGGGA AGGGGCAAGA
20351 GGATAAGTCA TGTGTCAGGA AGCAGCGTCC AACCAGAGCC GGCCACCTGT

FIGURE 3F

20401	CCCTTTTCT	GCCACCATGC	ACCAACTTTG	CTGTTTCAGTC	ACTGAAGCTC
20451	ATTCTGCAT	GGCTTCCTCC	CTTCCAGGCT	CCAGGGGATG	CTGAGCTCTC
20501	CTTGACGAC	TCAGAGGAGG	ACATTTCGAG	CTACTGGGAC	TTAGAGCAGG
20551	TGAGCTGAGG	GAAGGGGCTG	TGAGGGTGGG	AGCAGGGCGA	AGAGGGGAAG
20601	GATGGGGTCG	CTGTCAAATA	CAAGGCGTTC	ACTCAGCTGT	CTCACCTCCA
20651	GCCCAGAGCA	GTCACATTCA	AGGCCACAAA	GATTTGTGGT	CATCTTTGTT
20701	TTTTTCTTT	TCCTTTTCTT	TTTTTTTTTT	TTTTAATTTG	AGACAAAGTC
20751	TCACCTATC	ACCCAGACTG	GAATGCAGTG	GCATGATCTC	AGCTCACTGC
20801	AACCTCTGCC	TCCCGGGTTC	CAGAGGTTCT	CCTGCCTCAG	CCTCCCGAGT
20851	AGCTGGGACT	TCAGGCCTGC	GCCCAGCTAA	TTTTTGATT	TTTAGTAGAG
20901	ACAGCTTTTC	ACCATGTTGG	CTGGGCTGGT	CTCGAACTTC	CGATCTCAAG
20951	CAATCTGCCT	GCCTCGGTCT	CCTAAGTGCC	TGGATTACAG	GCATAAGCCA
21001	CGATGCCTGG	CCTTTGTTTT	CATTCTTCTC	ACTCCCTGAA	AGGCATCGTG
21051	GGGAGAGGGT	GAGTCACTGG	ACCAAGTCCT	AGAGAACCAG	TATCTATTCT
21101	TATTCTCCAA	CACATCACCC	ACGTGACCCT	GAGCAAGCCA	CATACACCCT
21151	GGGCCCTAGT	TTTTATCATC	TGTGAAATTA	GGGGAAACAT	AGGTAATACC
21201	TGTCCCATCC	ACCACACAAG	ATTGGCAGGG	CAGTCACTTG	TTCTTTTATT
21251	AATTACAGCAG	GTATTTATGG	CGTACCTACT	GTTTGCCTGA	CACAGTTTCA
21301	GATGGGCACA	TAGCAGTGAG	CAAAACAAAG	GCCTCTGCCT	TTTAGAAACT
21351	TACGTTATGG	TAGAATAGAT	GGATTTNNNN	NNNNNNNNNN	NNNNNNNNNN
21401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNGTCT	ACAAATGAAT	TATTATTGCA
21451	TGTGGACAAG	AGTTAAGAAC	TAAAAAATAT	GTGGCTGGGT	GCAATGGTTC
21501	ACACCTGTAA	TCCCAGCACT	TTGGGAGGCT	GAGGTGGGCG	GACCACCTGA
21551	GGTCAGGAGT	TTGAGACCAG	CCTGGCCAAC	ATGGCGAAAC	CCCGTCTCTA
21601	CTAAAAGCAC	AAAAATTAGC	CAGGCGTAGT	GGTGCATGCC	TGTAGTCCCA
21651	GCTACTCGGA	AGTCTGAGGC	ATGAGAATCA	CTTGAACCTG	GGAGGCAGAT
21701	GTTGCAGTGA	GCCGAGATCG	TGCCACTGCA	CTCCAGCTTG	GGTGACAGAG
21751	CTAGACTGTC	TCAAAAACAA	ACAAACAAAA	CAAAACCTAA	AAGATATGTG
21801	GATATGAGGG	ATCACCATCC	CCATAGGGCC	CCTGGATTAA	CACCACCCCA
21851	CCAATGCCCT	GAATTAAGAG	AAACCAGATG	ACTAGGTTTG	GAGAAATCTG
21901	GCTTTGGGTC	TATGAGAAAGT	AGTGTCTCTC	TTTGTGCCTC	TTCCCATTTCT
21951	TTTTGACATT	GAGCTCCATG	GTGCTCTGAA	TCCGTCTCTC	ACAGTGCTGA
22001	TGGCAGGTGG	GACAGATTAG	AAAATAGAGC	TGGAGCCACA	GAGATTTGGC
22051	AGACTGATTT	CGGTGCCCTC	TTGGAATCTC	CAGCACATTC	CAAAAAGCCT
22101	GGATAGGACC	AAAATAGCTT	ATCAACGTGA	GAAAGGACTT	CAGAGCTTGT
22151	CTACTGCCAA	CCCTCATTTT	ACCCAATGAG	GAAAGTGAAG	CTATTAGGGG
22201	GCGAGGGACA	CGTGAAGGCT	CACACAGCAG	ACAGGAGGTG	ATTACATGTT
22251	AGATTTTCA	ACCTGCTCCT	GCCACGCTGG	ACTGGTTTCA	CTCCTAGGCT
22301	GACCCTGCCT	CTCCCCTGTT	CACACACACT	CTCGCACACA	CACACACACA
22351	CACACACACA	CACAGGTGCT	TTGTTCTGGC	CAGGGGTTCC	TAGGGTCACC
22401	TCTTGGTTGC	AGCCACTGTG	ACCCCAACTG	GTCTAACCTC	TCTCTTCCCC
22451	TCCCACCTTC	TTCTGTGGT	TCCTGCAGGA	GATGTTTCGGG	AGCATGTTTC
22501	ACGCAGAGAC	CCTGACCGCC	CTGTGAGGGC	TCAGCCAGTC	CTCATGCTGC
22551	CTACAGAGTG	CCTGGCACTT	GGGACTTCCA	TAAAGGATGA	GCCTGGGGTC
22601	ACAGGGGGTG	TCGGGCGGAG	GAAAGTGCAT	CCCCAGAGC	TTGGGTTTCT
22651	CTCTCCCTC	CCCTCTCTC	CTCCCTTCTC	TCCCTCCCCG	CATCTCCAGA
22701	GAGAGCCTCT	CAGCAGCAGG	GGGGTGCTAC	CCTTACAGGA	GTGAGAGTCT
22751	GGTGAGCCCA	CTCTTACCCC	GTGAGGCTCT	GGCCGCAATG	GACAAGCCTC
22801	CTGCTCACTC	CACCCACCCC	ACCTCTGCCC	TGTCCTTGGC	AGCTGAAGGA
22851	CACCTTGACT	TCCAGCTTTT	ACGAGTGAGC	CAAAAACAGA	AGGACAAGTA
22901	CAACTGTGCT	GGCCTGCTGT	ACAAGCTTCA	AAAAGTGTCC	CAGAGCCAC
22951	ACGGCTCGGT	GTCAGATGGT	GTGAGGCTGT	CACGGACATA	GGGATAAACT
23001	TGGTTAGGAC	TCTGGCTTGC	CTTCCCCAGC	TGCCTCAACT	CTGTCTCTGG
23051	CAGCTCTGCA	CCCAGGGACC	ATGTGCTCTC	CACACCCAGG	AGTCTAGGCC
23101	TTGGTAACTA	TGCGCCCCCC	GTCCATCATC	CCCAAGGCTG	CCCAAACCAC
23151	CACTGCTGTC	AGCAAGCACA	TCAGACTCTA	GCCTGGACAG	TGGCCAGGAC
23201	CGTCGAGACC	ACCAGAGCTA	CCTCCCCGGG	GACAGCCAC	TAAGGTTCTG
23251	CCTCAGCCTC	CTGAAACATC	ACTGCCCTCA	GAGGCTGCTC	CCTTCCCCTG
23301	GAGGCTGGCT	AGAAACCCCA	AAGAGGGGGA	TGGGTAGCTG	GCAGAAATCAT
23351	CTGGCATCCT	AGTAATAGAT	ACCAGTTATT	CTGCACAAAA	CTTTTGGGAA
23401	TTCTCTTTTG	CACCCAGAGA	CTCAGAGGGG	AAGAGGGTGC	TAGTACCAAC
23451	ACAGGGAAAA	CGGATGGGAC	CTGGGCCCCAG	ACAGTCCCCC	TTGACCCAG
23501	GGCCCATCAG	GGAAATGCCT	CCCTTTGGTA	AATCTGCCTT	ATCCTTCTTT
23551	ACCTGGCAAA	GAGCCAATCA	TGTTAACTCT	TCCTTATCAG	CCTGTGGCCC
23601	AGAGACACAA	TGGGTCCTT	CTGTAGGCAA	AGGTGGAAGT	CCTCCAGGGA
23651	TCCGCTACAT	CCCCTAACTG	CATGCAGATG	TGGAAAGGGG	CTGATCCAGA
23701	TTGGGTCTTC	CTGCACAGGA	AGACTCTTTA	ACACCCTTAG	GACCTCAGGC
23751	CATCTTCTCC	TATGAAGATG	AAAATAGGGG	TTAAGTTTTT	CATATGTACA

FIGURE 3G

23801 AGGAGGTATT GAGAGGAACC CTA CTGTTGA CTTGAAAATA AATAGGTTCC
23851 ATGTGTAAAGT GTTTTGTAAT ATTTCAAGTGG AAATGCACAG AAAATCTTCT
23901 GGCCTCTCAT CACTGCTTTT CTCAAGCTTC TTCAGCTTAA CAACCCCTTC
23951 CCTAACAGGT TGGGCTGGCC CAGCCTAGGA AAACATCCCC ATTTCTAACT
24001 TCAGCCAGAC CTGCGTTGTG TGTCTGTGTG TTGAGTGAGC TGGTCAGCTA
24051 ACAAGTCTTC TTAGAGTTAA AGGAGGGGGT GCTGGCCAAG AGCCAACACA
24101 TTCTTGGCCC AGGAGCATTG CTTTCTGTG AATTCATTAT GCCATCTGGC
24151 TGCCAATGGA ACTCAAACT TGAAGGCGA AGGACAATGT TATCTGGGAT
24201 TCACCGTGCA CAGCACCCGA AGTGCCAAAT TCCAGGAGGA CAAGAGCCTT
24251 AGCCAATGAC AACTCACTCT CCCCTACTCC ACCTCCTTCC AAGTCCAGCT
24301 CAGGCCCAGG AGGTGGGAGA AGGTCACAGA GCCTCAGGAA TTTCCAAGTC
24351 AGAGTCCCCT TTGAACCAAG TATCTAGATC CCCTGAGGAC TTGATGAAGT
24401 GATCCTTAAC CCCCAAGTAA TCATTAACCC CCAGACCAGC CTCAGAACTG
24451 AAGGAGATTG TTGACCAGT GACCTGGAGT TGAGGCTCAG GGAGAGATCT
24501 GCCACATGTC TGAGGGTTGC AGAGCC (SEQ ID NO:3)

FEATURES:

Start: 1997
Exon: 1997-2121
Intron: 2122-4732
Exon: 4733-4872
Intron: 4873-5004
Exon: 5005-5115
Intron: 5116-5781
Exon: 5782-5957
Intron: 5958-7770
Exon: 7771-7935
Intron: 7936-8470
Exon: 8471-8623
Intron: 8624-8917
Exon: 8918-9000
Intron: 9001-9777
Exon: 9778-9925
Intron: 9926-10221
Exon: 10222-10335
Intron: 10336-10539
Exon: 10540-10617
Intron: 10618-11197
Exon: 11198-11293
Intron: 11294-13338
Exon: 13339-13445
Intron: 13446-14214
Exon: 14215-14284
Intron: 14285-14400
Exon: 14401-14493
Intron: 14494-15980
Exon: 15981-16262
Intron: 16263-17597
Exon: 17598-17652
Intron: 17653-18842
Exon: 18843-18988
Intron: 18989-20477
Exon: 20478-20549
Intron: 20550-22478
Exon: 22479-22523
Stop: 22524

CHROMOSOME MAP POSITION: Chromosome 1

ALLELIC VARIANTS (SNPs):

DNA				Protein		
Position	Major	Minor	Domain	Position	Major	Minor
48	C	G	Beyond ORF(5')			
132	G	A	Beyond ORF(5')			
724	A	C	Beyond ORF(5')			
1558	C	G	Beyond ORF(5')			

FIGURE 3H

1577	A	G	Beyond ORF(5')			
2487	C	A	Intron			
2634	T	C	Intron			
4352	A	G	Intron			
5157	A	C	Intron			
5658	A	T	Intron			
5945	T	C	Exon	180	T	T
6281	C	T	Intron			
6452	G	C	Intron			
6610	T	G	Intron			
7247	T	C	Intron			
7360	A	G	Intron			
7644	A	T	Intron			
8127	A	C	Intron			
8317	G	A	Intron			
9079	G	A	Intron			
9537	G	T	Intron			
12302	C	G	Intron			
12354	C	T	Intron			
12487	C	T	Intron			
13198	-	A	Intron			
13257	A	G	Intron			
14541	G	A	Intron			
14545	A	G	Intron			
15041	C	A	Intron			
15053	A	C	Intron			
15065	A	G	Intron			
15108	A	C	Intron			
16274	-	G	Intron			
17424	C	T	Intron			
17627	G	A	Exon	657	V	V
18427	T	C	Intron			
18813	C	G	Intron			
19035	T	C	Intron			
19182	T	C	Intron			
19508	-	G C	Intron			
19571	T	G C	Intron			
20147	T	G	Intron			
20180	G	A	Intron			
20584	A	T	Intron			
20717	T	C	Intron			
20894	A	G	Intron			
21787	-	A C	Intron			
22264	T	C	Intron			
22338	-	C A	Intron			
23363	T	C	Beyond ORF(3')			
23688	G	A	Beyond ORF(3')			
24210	A	C	Beyond ORF(3')			

Context:

DNA

Position

48

CTGGGTTCTATGTGGGGAGGTCATGCTCCCCACTCATTGAGCCCCC
[C, G]

CAGGCAAACCACTGGACAGCCAGACCCATGCAGACTCTGGAGCAGGTGGAGAGGAAGAG
TGAGACCACCCCGCTCACGGGCGGTGAAGGGCCGGCAGCCTCTGAATAGTCTCTGCTAG
GAGGTAGAAAGCACCTCCCATCTTAATCATAGTAATCATCGCCACTACCATTCTACTGGG
TGCCTATAAAAGGCCAGCCTCTTCATACACATGATCTCACTGAATCCTCATAGCATCTGC
CTGCGACTGTTATTATCCCCATTTACAGATGAAGAACTGAATCTTTGAACCCAGGTCAT

132

CTGGGTTCTATGTGGGGAGGTCATGCTCCCCACTCATTGAGCCCCCCCAGGCAAACCACT
CTGGACAGCCAGACCCATGCAGACTCTGGAGCAGGTGGAGAGGAAGAGTGAGACCACCCC
GCCTCACGGGC

[G, A]

GTGAAGGGCCGGCAGCCTCTGAATAGTCTCTGCTAGGAGGTAGAAAGCACCTCCCATCT
TAATCATAGTAATCATCGCCACTACCATTCTACTGGGTGCCTATAAAAGGCCAGCCTCTTC
ATACACATGATCTCACTGAATCCTCATAGCATCTGCCTGCGACTGTTATTATCCCCATTT

FIGURE 3I

ACAGATGAAGAACTGAATCTTTGAACCCAGGTCATCTGGCTCTCAAACCTTGCTGCTTT
TCCCTAAGCCACCCGGTCTCTCATTTCTCCACTGAAATGTCTCACATGCCATTGCCCTT

724 ATTGCCCTTACTCATTTCTGCCATGTCTCTCCAAAACACCATTTATCAATTCGCTCAA
CAAGTATGTGTTGAGTACACACTAAGGGCCAGGCGAGGGCTGGGCACAGGCGCTGGGG
TAGGTTCAATCTCCACCTTCGCTTCTGCTGGGTATCACCTGTGGGGTCTTGCCGGGCAT
CCCACCTCACCTGTAGTTCAAGTGACCTTGGGATCCCAAGACCAATGAATGGAATGC
ACCAGCCCAGCCTTACCAACTTGAGCACAATCTTATTATAATAGAACTCACATTTGC
[A, C]
TCACACTTTACATTTTACACAACCCCTTCTTATCCATTAACCTCATTTGATCTTCACAACA
ACCCTGTGAGATATGTCTGTTACTCCACTTTAGTGATACAGAATCTGAGGTTTGAAAAG
TAATGCTGACCATTCTGCCTCATTAATAAAAGCAGGATTAACCCAGGCTCCTGGACCTT
CCACAAAAGGCATTAAGCAACCTGCTCCCCTCTGACAACCTCCCCTGTACCCAGGCTCT
CCTCTGGGAAGTTGGGGGCATCTCTAGCCCCAAGTAGTTACTCATTTTCAACCCCATCT

1558 TCAGCTCTGCCATCTCAGCTCCTGGAACGTGAGCCAGGTTGCGCAAAAAGTGAGGAGGA
GAGGAGCGGCAGTACACAAGGGTGGGGGAAAGATTAGGCACAGGAAGCCGTGGGAGAGAG
AGCCGGCAGGTGGACCATCTGGTTTCCCCACACACCACTTGTCCCCTGGGAAACCTG
TTGCTGAAGTTCTAGATGTCTTATCCAAGAAGGGTCTCTTGAGGTCTCTCAGCTATCC
CCCTGCCTCTAGGCAAGCTGTTTTCTGTTTCTTCAAGCTGACTGGCTGAATGGTAGGAG
[C, G]
CTTTCTGCCAGGGAACTAAGGTCTGGGAAGGGAGTATGGCTTGTTGGGGACACCAGGGGT
CAGGGGAGGGGAGGGTCCACCTGCTGAATCAAGTGGGGCTCTGCCCCTCGTGATTTCCC
TTTGCTGCTGCTCAGTGGGGGTGATGGTGACGCCACAGGTGTGGAGTGCCAGCCACGTG
CTGAGCGCCAAGCAAAACAGCCAGGGTGAGTCTATGCATCATCAGTGCCTGGGAAGGAAG
GCCACTGCGAGCAGGGAGTCTGACGGAAAACTTGACAGAGGGAAGGGAGGCACCTTGCT

1577 CTCCTGGAACGTGAGCCAGGTTGCGCAAAAAGTGAGGAGGAGAGGAGCGGCAGTACACAA
GGGTGGGGGAAAGATTAGGCACAGGAAGCCGTGGGAGAGAGAGCCGGCAGGTGGACCATC
CTGGTTTCCCCACACACCACTTGTCCCCTGGGAAACCTGTTGGTGAAGTTCTAGATGT
CTTATCCAAGAAGGGTCTCTTGAGGTCTCTCAGCTATCCCCTGCCTCTAGGCAAGCT
GTTTTCTGTTTCTTCAAGCTGACTGGCTGAATGGTAGGAGCCTTTCTGCCAGGGAACT
[A, G]
AGGTCTGGGAAGGGAGTATGGCTTGTTGGGGACACCAGGGGTGAGGGGAGGGGAGGGTCCA
CCTGCTGAATCAAGTGGGGCTCCTGCCCTCGTGATTTCCCCTTTGCCTGGTGCTCAGTGG
GGGTGATGGTGACGCCACAGGTGTGGAGTGCCAGCCACGTGCTGAGCGCCAAGCAAAACA
GCCAGGGTGAGTCTATGCATCATCAGTGCCTGGGAAGGAAGGCCACTGCGAGCAGGGAGT
CTGACGGAAAACTTGACAGAGGGAAGGGAGGCACCTTGCTTTATCGGGGCGGGGAAGGC

2487 ACACGGCTTCTGCACTGGTATCCCTAAGATGGGGTTAAGGGAAGCCCTGGGGAAGTGAGG
TTCTGAATGATGAATTTAAGATCCTACAACCTCATCTGTACTGAGACCCCAAGGAGGAT
GGGGAGCAGGAGCAAGAACCATCCAGAAGGGTTATATGGCATTCCCAAACCCCTGCATGG
CATCTCCATATTCTCAATTACCCGGGTCTCTCTGGGTTTGTAAAGGCATGGTAGATGA
GCATCTACGTTATGGAGGGGTGGGGAGCATCAGAGCCCTTACTCCATGCCCTGTTCCCTC
[C, A]
TTACAAAAAATACCTGAAGTTACCATCACCCAGGTTCTTTGTCTTTCCCTCCCGGATG
TTCCTTCTCCACTTGGTCCAGAGAATGCCAAAAGGAGGCCCTAAATTTCTGAACCTTCC
TGAGGGGACCTACCAGGGTGTAGTCTACCAGCGCCAGGGTCTTTCCACTCTCATCTCC
CTGGAAATGCGATGGTGGGTATGAAACCTTGTCCCTAAGTAGGCGCTACACAAGGTGATC
CATACCCACACCCAGGAGGCTGGGGCTGCGGGTGTCAACCTCCCCATTCCCAGACTCCT

2634 AGGGTTATATGGCATTCCCAAACCCCTGCATGGCATCTCCCATATTCTCAATTACCCGG
GTCTCTCTGGGTTTGTAAAGGCATGGTAGATGAGCATCTACGTTATGGAGGGGTGGGGAG
CATCAGAGCCCTTACTCCATGCCCTGTTCCCTCCTTACAAAAAATACCTGAAGTTACCAT
CACCCAGGTTCTTTGTCTTTCCCTCCCGATGTTCTTCTCCACTTGGTCCAGAGAA
TGCCAAAAGGAGGCCCTAAATTTCTGAACCTTCTGAGGGGACCTACCAGGGTGTAGTCC
[T, C]
ACCAGCGCCAGGGTCTTTCCACTCTCATCTCCCTGGAATGCGATGGTGGGTATGAAAC
CTTGTCCCTAAGTAGGCGCTACACAAGGTGATCCATACCCACACCCAGGAGGCTGGGGC
TGCGGGTGTCAACCTCCCCATTCCAGACTCCTGGCAGACCTCCTCTGGCCAGCTATAG
GCCAACTCACTCTCCCTCACTCCCTTGGGGAAACGGCTGATTCAGTTACCTGGATTGAGG
TCACTGGCAATGGCTGAAGTGAGACGCAGGTGGAACCTGGTTCAGGCCGGGGGAATCACC

4352 ATTGAGGTTACCACACATAAAGGATAGTGAGTCAGCAGAGTGCACCCTGCAGGAACAATA
GAGCCTTCTTTTTCAAGGAAGTTCTAAGAAAAATGGCAGCAGGCAGGCCCACTCGGGTG
TATTCATCATTCATTTATTCAACAAATATTTACTAAGTGCCCCCTGTGCAAGGCTCGAGG
TGACAAAGATGAACAGGAGAGCTAGACTTCTGCCATGCGTGGTGGGGTTTGTGCTCCTA
GTGGGAGAGACAGACAAAAAGCAAGGAATGCACACACAGGATGCACACACAGCGGCAGGA

FIGURE 3J

[A, G]
CCAAGGTGCAGTTACCCAGGCCTGGGATCAGACAGACAGGACTCAGAGGAGACTTTCCCA
GAGAAAAGCCACTCTGAGCCAAGGGATGGATCTGATACCTCCGAAGGCTGAGCCACCATAA
CACTCATACCTTTAAGCCAAGTCTTATAAACTCCCAGGTAAGCAGCTGGCAGTCAGAAG
ACCTCCAGCTAATGCCCAGGACAAGTTGATGAGCTCTCAAGAAAAAGTTCTGCCTTTTC
TTCTCAATATCCCTGGCACACAGTTCAAGTGAATTTTGAATGAACCAATGAATGAAATGAG

5157 ATCCAGGTCCCACAAGGTGAAGGGGCTCCTTCAGCCAGGCCTGGATTGCCACTCCCCTCA
CCATTCTCTCCTCATCCCCACTCCCTCTGTGATCCCATAAAGCTAGTCATGCTGC
TGAGCTTCAGTCTCGTTGTCTCTGCAGGCATGGCATTGCTCTGCTGGCCAACCTTCT
GCAGTCAATGGCCTCTACTCTCTCTTCCCCCTCTGACCTACTTCTCTGGGGGGT
GTTCAACAGATGGTGCCAGGTAAGGCCTCTCCCCTCTGGGCAGGCAGGATGACCCAGACC
[A, C]
CAAGGATGGGAGGTGTGGCAAAGGGGCTCGGGAGATTTTCCATCTGCATTCTCTGGAG
TTGTTCTGGTCAGTCTAGGGGAATGGTCACTGTGAATGTCAATTCAGGTCTCTGGTG
ACCTTGAGAAACCACTGAGCCTCTTTGAGTTCAGTTAGCATTACCTGTTCCATCTTCT
CCTAGGAATGAGAGGAAGACTTAGCAGAAACAAGATATACCATATGCTATAACATGCTTAA
ACAGATGTGAGAAATCACCATCTAACTCCCTGGTTGGTCCCAGCCGGCCACTACAGGGAG

5658 TTAGCAGAAACAAGATATACCATATGCTATAACATGCTTAAACAGATGTGAGAAATCACCA
TCTAACTCCCTGGTTGGTCCCAGCCGGCCACTACAGGGACATTTGGACTTCTCTGGTGCT
AAGTGAGATGGAGGAAAGCCTGGTCACAAGGGCTGGTTTCTGGTTCAAGCTCTGCTTATA
TTTCTTATTTCTGAGTTCAATTTCTCAGGTGCTGTATGACAATATTGACCATTTGGGT
AAAAGCACCTTGAAGAACATAGATCATGGTTAGAGTGAGTGGTTGTTATTATTGTTGG
[A, T]
GAAGAGCCTTGAGGTGCAGGGATCCATCCCCCTGGGGTGGGAAGCATTCTCTGGGCC
TTCTGGTTTCCATCGGTGTGGTTCAAACCTCTGATTTTTGCTGGCTGGGTGGGGACCA
CAGGTACCTTTGCCGTTATCAGCATCCTGGTGGGTAACATCTGTCTGCAGCTGGCCCAAG
AGTCGAAATTCAGGTCTTCAACAATGCCACCAATGAGAGCTATGTGGACACAGCAGCCA
TGGAGGCTGAGAGGCTGCACGTGTGAGTACGCTAGCCTGCCTCACTGCCATCATCCAGG

5945 ATTATTGTGTTGGAGAAGAGCCTTGAGGTGCAGGGATCCATCCCCCTGGGGTGGGAAG
CATTCTCTGGGCCCTTTCTGGTTCCATCGGTGTGGTTCAAACCTCTGATTTTTGCTGGC
TGGGTGGGGCACCACAGGTACCTTTGCCGTTATCAGCATCCTGGTGGGTAACATCTGTCT
GCAGCTGGCCCCAGAGTCGAAATTCAGGTCTTCAACAATGCCACCAATGAGAGCTATGT
GGACACAGCAGCCATGGAGGCTGAGAGGCTGCACGTGTGAGTACGCTAGCCTGCCTCAGC
[T, C]
GCCATCATCCAGGTGAGGGGGCAGCCCCCAACCCTGCTAGAAGGGCATCAGACCACCCTG
CCCCTCCCTCAAAGCCTTAGCTTTGATGCTAAATCTGATTTAGGGGGCTGGGTGTGGAG
CTCATGCTGTAATCCAGCACTTTGGGAGGCTGAGGAGGGTGGATCACTTGAGGTGAGG
AGTTTGAGACCACCTTGACCAACGTGATGAAACCCCATCTCTACCAAAAAATACAAAAATA
ATCCAGGCTTGTTAGTATGCGCCTGTAGTCCCACCTACTCAGGAGGCTGAGGCAGGAGAA

6281 GCTAGAAGGGCATCAGACCACCTGCCCTCCCTCAAAGCCTTAGCTTTGATGCTAAATC
TGATTTAGGGGGCTGGGTGTGGAGGCTCATGCCTGTAATCCAGCACTTTGGGAGGCTGA
GGAGGGTGGATCACTTGAGGTGAGGAGTTTGAGACCACCTTGACCAACGTGATGAAACCC
CATCTTACCAAAAAATACAAAAATAATCCAGGCTTGGTAGTATGCGCCTGTAGTCCCACC
TACTCAGGAGGCTGAGGCAGGAGAACTCACTGAATCCGGGAGGCAGAGGTTGCAGTGAGC
[C, T]
GAGATCGCGCCACTGCACTCCAGCCTGGGTGACAGAGCGAGACTCCGTCTCAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAGTTAGGGCTCACCTCCTCCTCCTCCCATCCCAGG
GCTAAAGTGAACCTTGAAAAATTAACAGTATCTCCTCATCTGCATGTAGCAGGACCATACA
AAAAACAACAGCTGTACCTGGTTAACTGTCTGAGCTTTAAACCTGTAAAAGACTCAC
AGCCTCTCTCCATTATCCCCTGGAGAAACCAACTCTCTGCCAGCATAGTCTTGAGACT

6452 ATGAAACCCCATCTCTACCAAAAAATACAAAAATAATCCAGGCTTGGTAGTATGCGCCTGT
AGTCCCACCTACTCAGGAGGCTGAGGCAGGAGAATCACTTGAATCCGGGAGGCAGAGGTT
GCAGTGAGCTGAGATCGCGCCACTGCACTCCAGCCTGGGTGACAGAGCGAGACTCCGTCT
CAAAAAAAAAAAAAAAAAAAAAAGTTAGGGCTCACCTCCTCCTCCTCCTCCTCCTCCTC
CCATCCCAGGGCTAAAGTGAACCTTGAAAAATTAACAGTATCTCCTCATCTGCATGTAGCA
[G, C]
GACCATACAAAAACAACAGCTGTACCTGGTTAACTGTCTGAGCTTTAAACCTGTAA
AAGACTCAGCCTCTCTCATTATCCCGTGGAGAAACCAACTCTCTGCCAGCATAGTC
TTGCAGACTGCTAATTTTCTCTAACATCCCTCACTCGCTCCAGCCTCCTCTGCTCAAG
CCACAGCAGCAGTTGCACAACATAAATTGAGCTTCTGCAATGGTTGCAAAGGATTCTGC
TAGGTTTTATGAAGGGAAGCAACAATGACAGAATGCAAGAGCAAAACACAGTCCCAGAG

6610 GTGACAGAGCGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAGTTA

FIGURE 3K

GGGCTCACCTCCTCCCTCCTCCCCATCCCAGGGCTAAAGTGAACCTTGAAAAATTAACAGT
ATCTCCTCATCTGCATGTAGCAGGACCATACAAAAAACAACAGCTGTACCTGGTTAAAC
TGCTCTGAGCTTTAAACCTGTAAGAGCTCACAGCCTCTCCATTATCCCGTGGAGAAA
CCCAACTCTCTGCCAGCATAGTCTTGCACTGCTAATTTTCTCTAACATCCCTCACTCC
[T, G]
CTCCAGCCTCCTCTGCTCCAAGCCACAGCAGCAGTTGCACAACATAAATTGAGCTTCTGC
AAATGGTTGCAAAGGATTCTGCTAGGTTTTATGAAGGGAAGCACACATGACAGAATGCA
AGAGCAAAACACAGTCCCAGAGAGCGCCTTTTCATTCACTCATTTCGGTTTTGTGCC
AAGAACTAGGCTAAACCTGGGATACAAAGATAAGTAAGAAAGAGGTCCAATTCACAAGT
TGCTCACAGCCACAGAGGAAGGAGCCATGTCAACAGATAAATTTGTATGCAGTGAGAT

7247 GACACAGAGCAGAGTCACGGAGGACCTCAAAGAGGAGGTGACACTCCACCTCTCTTAAAG
GATGAGAACTTAACCAGGAACAAGGTATACAGAGGATGGTCAGGCAGAAGGGAACAGTG
CCTAAAAACACTGAGGCCTGAGAGAGTGTGATCTGCGCAGGCAAAGTAAGGGGCTTGGTG
TGGCTGGAGGTAGAGGGCCAGAAGAGGATGGAAGTAGGCAGGAGCCAGACAATGAG
ATCTGGAGCTGTCTCTGACAGCGACTTTGGGTCTGATTGGCAGTTTATAAGGATCGTT
[T, C]
GGGCTACACAATGATGAGTGGGAGGTGGATTAGAATCAAGGCAGGGGACCTGTTGGGAGA
CTCTGCAGAGGCCAGGCAGGAATAATGCAGGCGAAGACCAGGTAGAGAAAGAGATGGGG
CTGGACTTGAAAAGAATGTTTTACCAGGAGCTTGGTGATAGACTGGATGTGGGAGGTAAG
GGAGGATGACTCTCAAGTTTTTGGTTGGGCAACCAGGTAAATGATGGTGTCTTTACTGA
GAGAGAAAACACTGGGGGAGGACTAGACTTATTTTACAGATAAGCCAAAGCCAGAGAGGT

7360 AACAGTGCCTAAAAACACTGAGGCCTGAGAGAGTGTGATCTGCGCAGGCAAAGTAAGGGG
CTTGGTGTGGCTGGAGGAGTGGAGGCCAGAAGAGGATGGAAGTAGGCAGGAGCCAGA
CAATGAGATCTGGGCTCTGTTCTCTGACAGCGACTTTGGGTCTGATTGGCAGTTTATAAG
GATCGTTTTGGGCTACACAATGATGAGTGGGAGGTGGATTAGAATCAAGGCAGGGGACCTG
TTGGGAGACTCTGCAGAGGCCAGGCAGGAATAATGCAGGCGAAGACCAGGTAGAGAAAAG
[A, G]
GATGGGGCTGGACTTGAAAAGAATGTTTTACCAGGAGCTTGGTGATAGACTGGATGTGGG
AGGTAAGGGAGGATGACTCTCAAGTTTTTGGTTGGGCAACCAGGTAAATGATGGTGTCT
TTACTGAGAGAGAAAACACTGGGGGAGGACTAGACTTATTTTACAGATAAGCCAAAGCCA
GAGAGGTGATGTGACAGAAAGGCCATGCTCTAAGGAGCTGAAGGTCTGATGGCAGCCA
TGTAGAGCACAGTGAAGGGCAGGTGAAGGTACAGATGGTCCAATTCCTCAAGCTACTG

7644 GACCAGGTAGAGAAAGAGATGGGGCTGGACTTGAAAAGAATGTTTTACCAGGAGCTTGGT
GATAGACTGGATGTGGGAGGTAAGGGAGGATGACTCTCAAGTTTTTGGTTGGGCAACCAG
GTTAATGATGGTGTCTTTACTGAGAGAGAAAACACTGGGGGAGGACTAGACTTATTTTA
CAGATAAGCCAAAGCCAGAGAGGTGATGTGACAGAAAGGCCATGCTCTAAGGAGCTGA
AGGTCTGATGGCAGCCATGTAGAGCACAGTGAAGGGCAGGTGAAGGTACAGATGGTCCA
[A, T]
TTCCCTCAAGCTACTGCTACGCTAGGACTGCACGGAGCTCCAGACCTGCGTGTGTGTGGG
GCGGGTCGTTGGAAGTCTGAACCACTTGGTCTTCCGCCACCAACCACCTTTCTCCTC
TCTCAGATGGGTCTGGGCTTCATGCAAGTTTGGCTTTGTGGCCATCTACCTCTCCGAGTCC
TTCATCCGGGGCTTCATGACGGCCGCCGCTGAGATCCTGATTTGGTGTCTCAAGTAC
ATCTTCGAGTACCATCCCCTCCTACAGGCCAGGGTCCATCGTCTTTGTGAGTCTG

8127 CATCCGGGGCTTCATGACGGCCGCCGCTGCAGATCCTGATTTGGTGCTCAAGTACAT
CTTCGGACTGACCATCCCCTCCTACAGGCCAGGGTCCATCGTCTTTGTGAGTCTGGG
GATGCACCCCTGCCATTGGAGCAAGGCTCCAGCAGACACATGAGGAGGATGTACTGTTTT
AAGATGTCGTGAGCTCCTCATTGCAAGGGCTGGCTTAGCTGTTGTTGAGAGAGGATTCTG
AGGGGGTTTTCTGTCTTGGGAGGTCAAAGTCATGACTCACAGAGGTTCTTGGTAGTTAAT
[A, C]
CCTGCAGAAAAGAGCTGTACATTCTCCGCCAGTTCCTCATTCTAGTGCCTCAACCCCTCC
CTGCCTGAAAAGTCTGCCTTATGTCTAATCTCCATCCCTCCTCCTCAGCCCAAACCTCT
TCTAAAGAAAAAGAAAGCAATTCCTTTCTAGCACAAAGTTCCTCATGTGCTTTTGGGAAA
GGGCGGTGGGCGACGGGACAGGGTTCCTGATCAGGGTTTTAATTCTGTCTTGGTGTGCCT
CCATTAGCTTTGATGGCATCCCTCCCTGGGTGAGACACCCAAAGGTGGGGTATTATGGG

8317 GAGCTCCTCATTGCAAGGGCTGGCTTAGCTGTTGTTGAGAGAGGATTCTGAGGGGGTTTC
TGTCTTGGGAGGGTCAAAGTCATGACTCACAGAGGTTCTTGGTAGTTAATACCTGCAGAA
AAGAGCTGTACATTCTCCGCCAGTTCCTCATTCTAGTGCCTCAACCCCTCCCTGCCTGGA
AAGTCTGCCTTATGTCTAATCTCCATCCCTCCTCCTCAGCCCAAACCTCTCTAAAGAA
AAAGAAAAGATTCCTTTCTAGCACAAAGTTCCTCATGTGCTTTTGGGAAAGGGCGGTGG
[G, A]
CGACGGGACAGGGTTCCTGATCAGGGTTTTAATTCTGTCTTGGTGTGCCTCCATTAGCTT
TGATGGCATCCCTCCCTGGGTGAGACACCCAAAGGTGGGGTATTATGGGAAAGAGGGGT
GGGAGCCTGTGAGCATGATGCTCTTTCCCCAGACCTTCATTGACATTTGCAAAAACCTC

FIGURE 3L

CCCCACACCAACATCGCCTCGCTCATCTTCGCTCTCATCAGCGGTGCCTTCCTGGTGCTG
GTGAAGGAGCTCAATGCTCGCTACATGCACAAGATTGCTTCCCCATCCCTACAGAGATG

9079 TTCTACTGCTCTAATAATTCCCCCTAAGGAGGCAGGGGAGTGGGATTAGGGTCCCCAGA
GAAAAGGGAGACTTGAGAGAGACGCTGCCCTGGCCCCACCTTAGGGCCAATCCCCATTC
TCCACTCTGGGGTTTGAGGTGGTGGTGGCAACAGCTATCTCCGGGGGCTGTAAGATGCC
CAAAAAGTATCAGATCGATCGTGGGAGAAATCCAACGCGGGTGAATCCAGGTGGCCCA
GAAGCCTGGCCACCCGCACCTCATGCCCACTAAGGCTGAGCTCGGAGAGGGAGACAA
[G,A]
ATGAACTCTATGAAAGTGCAGTCGAAACTGTATGACACTGACCATGTATGAATTATTACT
ATTACCGTTTCCTGAGAAGGGCCGCACAACAGCCAATGTAGGCTATTTTATGAGAAATG
AGTCTTAAGTCCCACTCCCCTTATAAATCTCATTCAACTGATGCTGTTAAACAAAGCC
TCTCTGAACAGCCGCTTGTGGCTCTTTCCTTGTCTAATGCATTGGTTCTTGTCCAT
GTAGAAAGGGAATATTAGGTTCAACCAGATTATGAAGCATCACTCTGTGCCAGGCAC

9537 AACTGATGCTGTTAAACAAAGCCTCTCTGAACAGCCGCTTGTGGCTCTTTCCTTGTCTC
TAATGCATTGGTTCTTGTCCATGTAGAAAGGGAACTATTAGGTTCAACCAGATTTCATGA
AGCATCCACTCTGTGCCAGGCACCATGCTGGGCCCTGGGAGGAGAGGGGTGACGCTTGTCT
CTGCAGGGTTGGAAACAGGCAAGGGAGGGGAAGACCACATAGCACCAAGGTCTAGGGGTCT
GTGGACTCTGTAGCATAACAGGGTTCAGAATCTGGGAGTTAACAACAGAGGCCCTACCACA
[G,T]
ACTGGCCCCGGGACCTTGGGCAAGTTAGGTTCTCTCAGCCTCAGTTTCTCTCTTGTAA
ACAGGAGTGATGGTCCCTACCTATGGGGTGGTGTGAGGATTGAGCTGGATGGGATAA
CTTAGGCAAAGATCCCGGCACACCATGGGGCCTGGCTGGTCCCTGTGGGCTGGTGAAGG
ACTTGGCTGCCCTCCCACTCACACCTTGGGTTCTGCCTCCTTCTGGCTCCTCGGCAG
GTTCCCCACCCGGTGTGCGCTGTGGTCTCACAGTGAAGGACATGATAGGCACAGCCTT

12302 AGCCCCACCATAACCTATGGGAGAGGATTTACTAACTTTCTTAACGGTGATGAAACCAA
GGCTCAGAATGGTTAAGTAAATTTGTCAAAGGCCACAGAGGTAGGGAGTGGTAGAGTCTGG
ATTAAACTCCAAGTCTGGACTCCAGACCTTAGGCTGTACTGTCTCATAGGGAAGGCA
GTCTCACCCACCTAGGGCAGAGAAGAAAATCTTAAAGCCAGAGAAGTGAGTGGCTCATC
TGTGGTCAACCCAGAGAGACAGTGATGAGGACAGGGAGAAAAATTATACCTCAGTTCACG
[C,G]
CCAAGGATCTGCTTTGACCATAACCCAACAAGCCCCGCTATGGTGGTATTTCTTAGGT
TCATATGGCGGCTTTTGTTCATTTGATCTTCACAGCAATTCTCTACAGGAATCTGGGC
AGATTTATTTCTTTAGAGGAATTTCCAGGTCTTAAATCTATAGGGGGCACTATCAAA
ACTTCACCCAATGTTGCCCCCTACCCACACAAAAACCAGGCCCCAGCCGATCAGAAAG
CACTGCTGAGCTCCTGTAGGGCCACGCAGCTCGCTGTGAGACAGAGAGAGGGAATCA

12354 GAAACCAAGGCTCAGAATGGTTAAGTAAATTGTCAAAGGCCACAGAGGTAGGGAGTGGTA
GAGTCTGGATTAAACTCCAAGTCTGGACTCCAGACCTTAGGCTGTACTGTCTCATAG
GGAAAGGCAGTCTACCCACCTAGGGCAGAGAAGAAAATCCTTAAAGCCAGAGAAGTGAGT
GGCTCATCTGTGGTCAACCCAGAGAGACAGTGATGAGGACAGGGAGAAAAATTATACCTCA
GTTCCAGCCCAAGGATCTGCTTTGACCATAACCCAACAAGCCCCGCTATGGTGGTATT
[C,T]
CCTTAGGTTCATATGGCGGCTTTTGTTCATTTGATCTTCACAGCAATTCTCTACAGGA
ATCTGGGCAGATTTATTTCTTTAGAGGAATTTCCAGGTCTTAAATCTATAGGGGGCAA
CTATCAAACTTCACCCAATGTTGCCCCCTACCCACACAAAAACCAGGCCCCAGCCGA
TCAGAAAGCACTGCTGAGCTCCTGTAGGGCCACGCAGCTCGCTGTGAGACAGAGAGAG
GGAATCACAATTTATTGATCACCTACTGAGCATCCATCACTAGGCTAGGACCGTCACATT

12487 ACCCACCTAGGGCAGAGAAGAAAATCCTTAAAGCCAGAGAAGTGAGTGGCTCATCTGTGG
TCACCCAGAGAGACAGTGATGAGGACAGGGAGAAAAATTATACCTCAGTTCACAGCCCAA
GGATCTGCTTTGACCATAACCCAACAAGCCCCGCTATGGTGGTATTTCTTAGGTTTCAT
ATGGCGGCTTTTGTTCATTTGATCTTCACAGCAATTCTCTACAGGAATCTGGGCAGAT
TTATTTCTTTAGAGGAATTTCCAGGTCTTAAATCTATAGGGGGCACTATCAAACTT
[C,T]
ACCCAATGTTGCCCCCTACCCACACAAAAACCAGGCCCCAGCCGATCAGAAAGCACTG
CTGAGCTCCTGTAGGGCCACGCAGCTCGCTGTGAGACAGAGAGAGGGAATCACATTT
ATTGATCACTACTGAGCATCCATCACTAGGCTAGGACCGTCACATTCTTAACTTTTGA
ATCCTTTGATAGGTAGGCATTATTATTCTCTTTTGTTCATAGCCATTAAAGAACA
AAATTTGGGGTGGTGTGCTGACTCACACCTGTGATCTAGCACTTAGGGGGCTGAGGC

13198 CTAAGTATTAGGAAGGTTAGGCGGGAGCAAACTTGGGTTCCAGGGTTTGGGCTCCAG
TGAGCTGATCTTGCCACTGCATACAGCCTGAGCAACAGAGCAAGACCCTGTGACTCCAA
AAACAAACAAACAAACATTTTGAACCCAAACAGATCTGACCCAAAGATGCATGCTCTTA
TAGATGCCACCTCCCTGTGTGCTGGGCTTCTACTAAAAACACAGACAAGATCAGGCAAC
CACAGTCAATCTAAGGGAAAGAGGAAAGTGAACCAAGCAAAATACATAAAATATTGC

FIGURE 3M

[- , A]
AAAAATGCTATTTAAAGAAAAAAGAGAAGAGAGGCTCTGAGGTTGTAATAACAGAGAAT
GGCCTTGGCTAATCCAGGAAGACTTCCTGAAAGAGGTTGTTTTTTCCCCAGGTCTGCTTT
TGACATCTCTCTTTTACAGTGCATCTGGGTAGTGAGCTTCCTCTCCTCCTTCTTCTCA
GCCTGCCCTATGGTGTGGCAGTGGGTGTCGCCTTCTCCGTCTGGTCTGTTCTTCCAGA
CTCAGTTGTAAGTGATAGCTTCCGCCCTCTAGGCCACAGTCGGTTCCTGGGCCAGCC

13257 GTGAGCTGATCTTGCCACTGCACTACAGCCTGAGCAACAGAGCAAGACCCTGTGACTCCA
AAAACAAACAAACAAACACATTTTGAACCCAAACAGATCTGACCCAAGATGCATGCTCTT
ATAGATGCCACCTCCCTGTGTGCTGGGGCTTCTACTAAAAACACAGACAAGATCAGGCAA
CCACAGTCAATCTAAGGGAAGAGGAAAGTGTAACCAAAGCACAAATACATAAAATATTG
CAAAAATGCTATTTAAAGAAAAAAGAGAAGAGAGGCTCTGAGGTTGTAATAACAGAGA
[A, G]
TGGCCTTGGCTAATCCAGGAAGACTTCCTGAAAGAGGTTGTTTTTTCCCCAGGTCTGCTT
TTGACATCTCTCTTTTACAGTGCATCTGGGTAGTGAGCTTCCTCTCCTCCTTCTTCTC
AGCCTGCCCTATGGTGTGGCAGTGGGTGTCGCCTTCTCCGTCTGGTCTGTTCTTCCAG
ACTCAGTTGTAAGTGATAGCTTCCGCCCTCTAGGCCACAGTCGGTTCCTGGGCCAGC
CCGCAAAGGGCTTCATGCCACGGCTGGCTTAGTCCACTGTACCTTCCACCTCTGGGCC

14541 TCATGGACACTGACATTTATGTGAATCCCAAGACCTATAATAGGGTAGGTAATTCAAGCT
TATGACCTCCTTCTTTTGTCTGCAACACCCCAAGAAGAGGTTGCTTTTTAAAGCCAATA
AAGACATTTCTGCAACTTGAGCTCAGTCTCCCTGTACAGGCCCAGGATATCCAGGGGAT
TAAATCATCAGTACTGCTCCCCTCTCTACTTTGCCAACTCAGAGATCTTCAGGCCAAAA
GGTCATCGCCAAGGTAAGGCTCAGTCCCTGGCGACCAGAGGCTCTGGACAGAGAGTGGCC
[G, A]
GAAAATGGAAGCAGAAGGGCGGTGGGAGCTGAGAATAGGCCACTCCATAGAGGGTGGAG
GTCAAGATTGCTGTTGGCTCTCTCCCTGCAGACAGGCATGGACCCCAGAAAAGTATTACT
AGCCAAGCAAAAATACCTCAAGAAGCAGGAGAAGCGGAGAATGAGGCCACACAACAGAG
GAGGTCTCTATTATGAAAACCAAGGTGAATGAAGGCCAGAAGCAGCCCCGTGCCCTGCT
CTCCTGCCCATTTCTGATACTGCCCCCTGTTACTCATGGTACCCTGGGGGCCCCGCTTCCC

14545 GGACACTGACATTTATGTGAATCCCAAGACCTATAATAGGGTAGGTAATTCAAGCTTATG
ACCTCCTTCTTTTGTCTGCAACACCCCAAGAAGAGGTTGCTTTTTAAAGCCAATAAAGA
CATTTCTGCAACTTGAGCTCAGTCTCCCTGTACAGGCCCAGGATATCCAGGGGATTAAA
ATCATCAGTACTGCTCCCCTCTCTACTTTGCCAACTCAGAGATCTTCAGGCCAAAAGGTC
ATCGCCAAGGTAAGGCTCAGTCCCTGGCGACCAGAGGCTCTGGACAGAGAGTGGCCGGAA
[A, G]
ATGGAAGCAGAAGGGCGGTGGGAGCTGAGAATAGGCCACTCCATAGAGGGTGGAGGTCA
AGATTGCTGTTGGCTCTCTCCCTGCAGACAGGCATGGACCCCAGAAAAGTATTACTAGCC
AAGCAAAAATACCTCAAGAAGCAGGAGAAGCGGAGAATGAGGCCACACAACAGAGGAGG
TCTCTATTATGAAAACCAAGGTGAATGAAGGCCAGAAGCAGCCCCGTGCCCTGCTCTCC
TGCCCATTTCTGATACTGCCCCCTGTTACTCATGGTACCCTGGGGGCCCCGCTTCCCACCC

15041 ACCAAGGTGAATGAAGGCCAGAAGCAGCCCCGTGCCCTGCTCTCCTGCCCATTTCTGATAC
TGCCCCCTGTTACTCATGGTACCCTGGGGGCCCCGCTTCCCACCCTGACAGGCAAAGACA
GAAAGTCTCTGGGAACACTGCCTGGTGGCCGCTGGGCATTTTTCTTCTTTTTTTCTTTT
TCTTTTTAGAGATGGAATTTTGTCTTGTCAACCCAGGCTTGAGTGCAATGGCGTTATCTT
GGCTCACTGCAACCTCACCTCTGGGGTTCAAGCGATTCTCTGCCTTAGCCTCCCAAGT
[C, A]
GCTGAGATTACAGGTGCCACCACACCCAGCTAATTTTTGTATTTTTAGTAGATATTGGGT
TTCACCATGTTGGCCAGGCTGGTGTCAAACCTCTGACCTCAGGTGATCCACCTACCTTAG
CCTTCCAAAGTGCTGGGATTACAAGCCTGAGCCACTGCGCCAGCCTGGGCATTTTTCTT
CTTGGATGAGGTGCTACCATCTCCAGGGAAGCCACTGAACCCCCAAGGCCCTTCTCCAT
TTTCTGGCTAAGATAGGACATGGCCCATGGACTTTTGAACAACCCAGAGGGGGAACAGCA

15053 GAAGGCCAGAAGCAGCCCCGTGCCCTGCTCTCCTGCCCATTTCTGATACTGCCCCCTGTTA
CTCATGGTACCCTGGGGGCCCCGCTTCCCACCCTGACAGGCAAAGACAGAAAGTCTCTGG
GAACTCTGCTGGTGGCCGCTGGGCATTTTTCTTCTTTTTTTCTTTTTCTTTTTAGAGA
TGGAATTTTGTCTTGTCAACCCAGGCTTGAGTGCAATGGCGTTATCTTGGCTCACTGCAA
CCTCCACCTCTGGGGTTCAAGCGATTCTCTGCCTTAGCCTCCCAAGTCGCTGAGATTAC
[A, C]
GGTGCCACCACACCCAGCTAATTTTTGTATTTTTAGTAGATATTGGGTTCACCATGTTG
GCCAGGCTGGTGTCAAACCTCTGACCTCAGGTGATCCACCTACCTTAGCCTTCCAAAGT
CTGGGATTACAAGCCTGAGCCACTGCGCCAGCCTGGGCATTTTTCTTCTTGGATGAGGT
GCTACCATCTCCAGGGAAGCCACTGAACCCCCAAGGCCCTTCTCCATTTTCTGGCTAAG
ATAGGACATGGCCCATGGACTTTTGAACAACCCAGAGGGGGAACAGCAGTGAATTTCTGT

15065 CAGCCCCGTGCCCTGCTCTCCTGCCCATTTCTGATACTGCCCCCTGTTACTCATGGTACC

FIGURE 3N

TGGGGGCCCCGCTTCCACCCCTGACAGGCAAAGACAGAAAGTCTCTGGGAACACTGCCTG
GTGGCCGCTGGGCATTTTTCTTCTTTTTTTCTTTTTCTTTTAGAGATGGAATTTTGCT
CTTGTCAACCCAGGCTTGAGTGCAATGGCGTTATCTTGGCTCACTGCAACCTCCACCTCTG
GGGTTCAAGCGATTCTCCTGCCTTAGCCTCCCAAGTCGCTGAGATTACAGGTGCCACCAC
[A, G]
CCCAGCTAATTTTTGTATTTTTAGTAGATATTGGGTTTACCATGTTGGCCAGGCTGGTG
TCAAACCTCTGACCTCAGGTGATCCACCTACCTTAGCCTTCCAAAGTGCTGGGATTACAA
GCCTGAGCCACTGCGCCAGCCTGGGCATTTTTCTTCTTGGATGAGGTGCTACCATCTCC
CAGGGAAGCCACTGAACCCCAAGGCCCTTCTCCATTTTCTGGCTAAGATAGGACATGGC
CCATGGACTTTTGAACAACCCAGAGGGGGAACAGCAGTGAATTTCTGGGGAACCCAGGC

15108 TGTACTCATGGTACCCTGGGGGCCCCGCTTCCACCCCTGACAGGCAAAGACAGAAAGTCT
TCTGGGAACACTGCCTGGTGGCCGCTGGGCATTTTTCTTCTTTTTTTCTTTTTCTTTTT
AGAGATGGAATTTTGCTCTTGTCAACCCAGGCTTGAGTGCAATGGCGTTATCTTGGCTCAC
TGCAACCTCAACCTCTGGGTTCAAGCGATTCTCCTGCCTTAGCCTCCCAAGTCGCTGAG
ATTACAGGTGCCACCACCCAGCTAATTTTTGTATTTTTAGTAGATATTGGGTTTACC
[A, C]
TGTTGGCCAGGCTGGTGTCAAACCTCTGACCTCAGGTGATCCACCTACCTTAGCCTTCCA
AAGTGCTGGGATTACAAGCCTGAGCCACTGCGCCAGCCTGGGCATTTTTCTTCTTGGAT
GAGGTGCTACCATCTCCAGGGAAGCCACTGAACCCCAAGGCCCTTCTCCATTTTCTGG
CTAAGATAGGACATGGCCCATGGACTTTTGAACAACCCAGAGGGGGAACAGCAGTGAATT
TCCTGGGGAACCCAGGCAGCCAGGGCTAGCAAGGCTGGGGTGGCCATGGCAGTAATCCT

16274 CTTCCAGACTGTCTCCCTGCAGGAGCTGCAGCAGGACTTTGAGAATGCGCCCCCACCAGA
CCCCAACAAACACAGACCCCGGCTAACGGCACCAGCGTGCTATATCACCTTCAGCCC
TGACAGCTCCTCACCTGCCAGAGTGAGCCACCAGCCTCCGCTGAGGCCCCCGGCGAGCC
CAGTGACATGCTGGCAGCGTCCACCCCTTCTGTCACCTTCCACACCTCATCCTGGACAT
GAGTGGAGTCAGCTTCTGAGACTTGATGGGCATCAAGGCCCTGGCCAAGGTGAGGCCCTC
[-, G]
GGGACAGCAAGCACCACCCACTCCACCCCTCCGCTCTGCTCTCCACATTCCCTTTCTG
GGAGCCCTCATTTAGGAAGCTGAGGGAGGAAGCTCACTGGGGAGACTAACAGCTCCTAG
GAATCCCTCCTTTCCCAGACGCCACCAGGTTGAGACATTCTCCACAGAGCAGGCCCAGA
CGGCCCATGACAATGAGTGGCGGGACAAGTCTACCAGAGTTTCAAGCCCTGTGCTCCA
ACACCCCCAGCAGTGGCCATCCCAAGTCCCTCTCAGCCATCAGGAACCCACCCAGGTTCT

17424 AACATGGTGAAACCCCGCTCTACTAAAAATACAAAAATTAGCCAGGTGTGGTGACGGGC
CCCTGTAGTCCCAGTACTCGGTAGGCTGAGGCAGAGAATTGCTTGAACCCAGGAGGCGG
AGGTTGCAAGTGAAGCAAGATCGCGCCACTGCACTCCAGCCTGGGCAACAGAGTGAACCTC
CATCTCAAAAGAAAAAAGAAAAATATCTAGCCCCAAGAAGGGCCATGGTGACTTT
AAGTGCCCGCCACGTTGGCAAAAGTCCATTTCCGCTCCACTTCCAGAGAAACCGTCAGC
[C, T]
AACACTCCAGGGAGAAGTGGTGTGCTTTGCTGCTATTTTTGTCTTTGGCTGCTGGGCTCT
CAGGTTTGTATTTTGTGCTTTCCCTCTGAAGTACGTTTGTGAATCACTTTTGAGA
CCCACTCAGAACATTCTTTCTTTTGCCTCCCTACCCCAACACACTTCTAGCTGAGCT
CCACCTATGGGAAGATCGGCGTGAAGGTCTTCTTGGTGAACATCCATGGTAAGAGAAAGA
GGACATTTAGGGACTGAAAGACTGGCAAGGAGTGTGGGGTAGGAACAGGTTGGTGGGGTG

17627 AATATCTAGCCCCACAAGAAGGGGCCATGGTGACTTTAAGTGCCCGCCACGTTGGCAAAA
GTCCATTTCCGCTCCACTTCCCAGAGAAACCGTCAGCCAACACTCCAGGGAGAAGTGGTG
TGCTTTGCTGCTATTTTTGTCTTTGGCTGCTGGGCTCTCAGGGTTGCTTATTTGTTTGGC
TTCCCTCTGAAGTACGTTTTGTGAATCACTTTTGAAGCCACTCAGAACATTCTTTCC
TTTTGCCTCCCTACCCCAACACACTTCTAGCTGAGCTCCACCTATGGGAAGATCGGCGT
[G, A]
AAGGTCTTCTTGGTGAACATCCATGGTAAGAGAAAGAGGACATTTAGGGACTGAAAGACT
GGCAAGGAGTGTGGGTAGGAACAGTTGGTGGGGTCTGAATAGTGAGGAGTTGGAAAC
GAGAGCACCAGCTATCCCCACAAGCTGCTGCCTGCTCATAAAAGCTTCAGGTACAAGT
CCAAAGAGACTGGTCAGATTGCATAAACATCCTAGGGGCCTTAGTGACAGAGTGGGGTG
AGGAGGTGATGGAGTTACAGAAGGACAGCTAGGATTCTAATCTACCCATAACTAATTTG

18427 GGGTGCATATACACAGCCTCAAGGACGTGGCCACAGGGCAGCAGACATTTACATGACTAG
CATGTACGCAAAGTGAGAGATGTGGGAGCAAGTGACACAGACACACAGGAGAATGTGA
AGGGGCACATACACACACCCAGCTCCCTGCACTGGGTGAGACCCCTCCAGCAGGGCTG
GCAGTCCCAAGCTCCGATGGCCAGTTTCGGGGAGAGAATCTGCAAGTGGCAATGACCTG
CTATGATATGTTCTGGAGTTAGAAGCAGTGGATTCTCCCAACCTCACTGGACACCCCT
[T, C]
AGGAAACCATCTCTAGGATTAAGAGTAATCCACACAACTTCCAATGCCACACATTGGAA
GTTGCTGGAAAGGTCTGGGAAACAAGGAAGGATGGGTCTTGGGGGATAGAATGGC
AGCGGCCTCTCAAGGATGGCTTAGGCTTTTCACTCGAATCACCACAAAGTACTGACTC

FIGURE 30

CCTAAATCAAACCTGCTTCCTTCTGCTCTGGGTTGAAACTTCAGCATCCTCAAGTTCATGT
TGCCCTCTGCCGTCCAGAACTGATATTGCACTGCCAATGCCATGGCCCTCAGATACAGCA

18813 AGAGGAAGGATGGGTCTTGGGGGATAGAACTGGCAGCGGCCTCTTCAAGGATGGCTTAG
GCTTTTCCACTCGAATCACCACAAAGTACTGACTCCCTAAATCAAACCTGCTTCCTTCTGC
TCTGGGTTGAACTTCAGCATCCTCAAGTTCATGTTGCCCTCTGCCGTCCAGAACTGATA
TTGCACTGCCAATGCCATGGCCCTCAGATACAGCAAGAGCTGGGACCTCAGGCCTCTCCC
ATCCCTGCTCTGGTCTCACTATCTTCCCCACCCCCAGCTCCAATCCACAATGGCTGTTAT
[C, G]
TTTCTGAAGGTGATCTTTTCTCCTTCTAGCCCAGGTGTACAATGACATTAGCCATGGAGG
CGTCTTTGAGGATGGGAGTCTAGAATGCAAGCACGTCTTCCCAGCATACATGACGCAGT
CCTCTTTGCCAGGCAAATGCTAGAGACGTGACCCAGGACACAACCTTCCAAGGGGTAAG
GTTCTTGACCTGGGAATCCTAGGCTCCAAGGCACTGAAATAGCAGGACCAAGAGGCAT
TATTAGAAAGAACACAGGAGAAGGTTTAAGTTCCAATATCAAGTCTGCCATTTCAGTTTT

19035 GGACCTCAGGCCTCTCCCATCCCTGCTCTGGTCTCACTATCTTCCCCACCCCCAGCTCCA
ATCCACAATGGCTGTTATCTTTCTGAAGGTGATCTTTTCTCCTTCTAGCCCAGGTGTACA
ATGACATTAGCCATGGAGGCGTCTTTGAGGATGGGAGTCTAGAATGCAAGCACGTCTTTC
CCAGCATACATGACGCAGTCTCTTTGCCAGGCAAATGCTAGAGACGTGACCCAGGAC
ACAACCTTCCAAGGGTAAGGTTCTTGACCTGGGGAATCCTAGGCTCCAAGGCACTGAAA
[T, C]
AGCAGGACCAAGAGGCATTATTAGAAAGAACACAGGAGAAGGTTTAAGTTCCAATATCAA
GTCTGCCATTTCAGTTTTCTGAATCTGTTTCTTATCTATAGAATGAGCACCATCACTA
ACATTACCTACCTCTCTGCATTTTTCTTTATTTTGTGTTAGGGTTAAATGATAATTACA
TCTTTTGTGTCACTTGAAAGCACTTTGTGTATTGTAATAATTCTTTATCAATATAAGTTT
TCTGTTGCACAAACACCCAAAGCATAGTAGAGCAGGCCCACTCTGCTGGCATCGTTC

19182 AGGATGGGAGTCTAGAATGCAAGCACGTCTTCCCAGCATACATGACGCAGTCTCTTTG
CCCAGGCAAATGCTAGAGACGTGACCCAGGACACAACCTTCCAAGGGGTAAGGTTCTTGC
ACCTGGGGAATCCTAGGCTCCAAGGCACTGAAATAGCAGGACCAAGAGGCATTATTAGAA
AGAACACAGGAGAAGGTTTAAGTTCCAATATCAAGTCTGCCATTTCAGTTTTCTGAATCT
GTTTCCTTATCTATAGAATGAGCACCATCACTAACATTACCTACCTCTCTGCATTTTTT
[T, C]
TTTTATTTGTTTTAGGGTTAAATGATAATTACATCTTTTGTGTCACTTGAAAGCACTTTG
TGTATTGTAATAATTCTTTATCAATATAAGTTTTCTGGTGCACAAACACCCAAAGCATA
GTAGAGCAGGCCCACTCTGCTGGCATCGTTCCCTGCTCCTCCTCATCTCTTTCTAAAGG
GGGCTTTCGGGAAGGGAGGGGAGGGGAGTAAGCCTACCCATTTTAACTTACCGGAGCTTA
GAGATTTCAAGGCTGGTGAGGGATAAAGAGATTGGGTCTGAGTTTTGTCTCAGCTTTTTGA

19508 TAATTACATCTTTTGTGTCACTTGAAAGCACTTTGTGTATTGTAATAATTCTTTATCAAT
ATAAGTTTTCTGGTGCACAAACACCCAAAGCATAGTAGAGCAGGCCCACTCTGCTGGCA
TCGTTCCCTGCCTCCTCCTCATCTCTTTCTAAAGGGGGCTTTCGGGAAGGGAGGGGAGGG
GAGTAAGCCTACCCATTTTAACTTACCGGAGCTTAGAGATTTCAAGGCTGGTGAGGGATAA
AGAGATTGGGTCTGAGTTTTGTCTCAGCTTTTTGACATTTAATTTACTAGCTCAGTAAGT
[- , G, C]
ATACAAATGGGATACAAATAACACCATCTAAAACCTCCAGAAGACTGGGGAGTCAGAAAAA
TCCTACCTCCTTGGGGTCCCTGCCAGATCCCAGTCATCTAGCCCTCAGGGTCCCCTCCC
CCAGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
AAACCTCCCCTGCCCTCTGTGGGTATAAGAAAGGCTTTTCTGGCCCTAGAGCAATGATT
TGCTCTTTGCCTTAAGAGACTGATGAAGGTGAAACCATCTGTTCTAAGTGCTGAAAGACT

19571 AGTTTTCTGGTGCACAAACACCCAAAGCATAGTAGAGCAGGCCCACTCTGCTGGCATCG
TTCCCTGCCTCCTCCTCATCTCTTTCTAAAGGGGGCTTTCGGGAAGGGAGGGGAGGGGAG
TAAGCCTACCCATTTTAACTTACCGGAGCTTAGAGATTTCAAGGCTGGTGAGGGATAAAGA
GATTGGGTCTGAGTTTTGTCTCAGCTTTTTGACATTTAATTTACTAGCTCAGTAAGTCAT
ACAAATGGGATACAAATAACACCATCTAAAACCTCCAGAAGACTGGGGAGTCAGAAAAATC
[T, G, C]
TACCTCCTTGGGGTCCCTGCCAGATCCCAGTCATCTAGCCCTCAGGGTCCCCTCCC
AGCTCAGCTCCTGCCCTGGCCTCCCAAGACTCTTGTGTGCCCCAGCCCTGGGTAAAAA
CCTCCCTGCCCTCTGTGGGTATAAGAAAGGCTTTTCTGGCCCTAGAGCAATGATTTGC
TCTTTGCCTTAAGAGACTGATGAAGGTGAAACCATCTGTTCTAAGTGCTGAAAGACTGCC
CAGGAACACACAGGGCGCTGGCTCCTGCCCTCCATGCCTAGAGGGAAACCTGGGGAAC

20147 GCCTAGAGGGAAACCTGGGGAACAACGGGCTTTCTGCTTCGTGAAATTTGTCCGCAG
AGCAAAGAGGGAGATTCTGGAGGAAGCTGCATTAGTTGTAGTGCCCTAATCATGTTTCAG
CTACTCTAGTTGGTATGTATACTTGATTAGTCATAGCACTTATAAATAATTTATATTTTA
TATAATATATACTTACATATTATAGACCATTACAGATACAAATCACACACATAAACACA
CACCTTTTCAACAGCATTTGTGAGGGACAAAGCAGGCAAAGTGAGGCTGGTTATCAGACTT

FIGURE 3P

[T, G]
AACAGATTAGAAAATATATTCCCAGGAGGACAGGAATCCCCAAGGTCAGGCAGCTAGCC
AATAGTTTTCTAAGCTGAGTAAACCTTCCCTGCCCTCTAACGGCCCAAAAGGAGGGAA
GACCGCGATACACACCTGTCTGGTATAAGGGGGAAGACCACAGCCGTGCTGTTTTGTGA
GGCAGGTAAGGGAAGGGCAAGAGGATAAGTCATGTGTGAGGAAGCAGCGTCCAACCAGA
GCCGGCCACCTGTCCCTTTCTGCCACCATGCACCACTTTGCTGTTCACTCACTGAAG

20180 TTCTGCTTCGTGAAATTTGTCCGACAGCAAAGAGGGAGATTCTGGAGGAAGCTGCATT
AGTTGTTAGTGCCCTAATCATGTTCACTACTCTAGTTGGTATGTATACTTGATTAGTCA
TAGCACTTATAAATAATTTATATTTTATATAATATATACTTACATATTATAGACCATTCA
CAGATACAAATCACACACATAAACACACACCTTTTCAACAGCATTGTGAGGGACAAAGCA
GGCAAAGTGAGGCTGGTTATCAGACTTTAACAGATTAGAAAATATATTCCCAGGAGGACA
[G, A]
GAATCCCCAAGGTCAGGCAGCTAGCCAATAGTTTTCTAAGCTGAGTAAACCTTCCCT
GCCTCTAACGGCCCAAAAGGAGGGGAAGACCGCGATACACACCTGTCTGGTATAAGGGG
AAGACCACAGCCGTGCTGTTTTGTGAGGCAGGTAAAGGAAGGGGCAAGAGGATAAGTCA
TGTGTGAGGAAGCAGCGTCCAACCAGAGCCGGCCACCTGTCCCTTTCTGCCACCATGC
ACCAACTTTGCTGTTCACTCACTGAAGCTCATTCTGCACTGGCTTCCTCCCTTCAGGCT

20584 TGTCTGGTATAAGGGGAAGACCACAGCCGTGCTGTTTTGTGAGGCAGGTAAAGGAAGG
GGCAAGAGGATAAGTCATGTGTGAGGAAGCAGCGTCCAACCAGAGCCGGCCACCTGTCCC
TTTTCTGCCACCATGCACCACTTTGCTGTTCACTCACTGAAGCTCATTCTGCACTGGC
TTCTCTCCCTCCAGGCTCCAGGGGATGCTGAGCTCTCCTGTACGACTCAGAGGAGGACA
TTGCGAGCTACTGGGACTTAGAGCAGGTGAGCTGAGGGAAGGGGCTGTGAGGGTGGGAGC
[A, T]
GGGCGAAGAGGGGAAGGATGGGGTGCCTGTCAAATACAAGGCGTTCACTCAGCTGTCTCA
CCTCCAGCCAGAGCAGTCACATTCAGGCCACAAAGATTTGTGGTCATCTTTGTTTTT
TTCTTTCTTTCTTTTTTTTTTTTTTTAATTTGAGACAAAGTCTCACTCTATCACCC
AGACTGGAATGCAGTGGCATGATCTCAGCTCACTGCAACCTCTGCCTCCCGGGTTCAGA
GGTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTTCAGGCCTGCGCCAGCTAATTTT

20717 ATGCACCAACTTTGCTGTTCACTCACTGAAGCTCATTCTGCACTGGCTTCCTCCCTTCCA
GGCTCCAGGGGATGCTGAGCTCTCCTGTACGACTCAGAGGAGGACATTCGACGCTACTG
GGACTTAGAGCAGGTGAGCTGAGGGGAAGGGGCTGTGAGGGTGGGAGCAGGGCGAAGAGGG
GAAGGATGGGGTGCCTGTCAAATACAAGGCGTTCACTCAGCTGTCTCACCTCCAGCCAG
AGCAGTCACATTCAGGCCACAAAGATTTGTGGTCATCTTTGTTTTTTCTTTCTCTTT
[T, C]
CTTTTTTTTTTTTTTTAATTTGAGACAAAGTCTCACTCTATCACCCAGACTGGAATGCA
GTGGCATGATCTCAGCTCACTGCAACCTCTGCCTCCCGGGTTCAGAGGTTCTCCTGCCT
CAGCCTCCCGAGTAGCTGGGACTTCAGGCCTGCGCCAGCTAATTTTTGTATTTTGTAGTA
GAGACAGCTTTTACCATGTTGGCTGGGCTGGTCTCGAACTTCCGATCTCAAGCAATCTG
CCTGCCTCGGTCTCCTAAGTGCTGGATTACAGGCATAAGCCACGATGCCTGGCCTTTGT

20894 GGGGAAGGATGGGGTGCCTGTCAAATACAAGGCGTTCACTCAGCTGTCTCACCTCCAGCC
CAGAGCAGTCACATTCAGGCCACAAAGATTTGTGGTCATCTTTGTTTTTTCTTTTCC
TTTTCTTTTTTTTTTTTTTTAATTTGAGACAAAGTCTCACTCTATCACCCAGACTGGAA
TGCAGTGGCATGATCTCAGCTCACTGCAACCTCTGCCTCCCGGGTTCAGAGGTTCTCCT
GCCTCAGCCTCCCGAGTAGCTGGGACTTCAGGCCTGCGCCAGCTAATTTTGTATTTT
[A, G]
GTAGAGACAGCTTTTACCATGTTGGCTGGGCTGGTCTCGAACTTCCGATCTCAAGCAAT
CTGCCTGCCTCGGTCTCCTAAGTGCTGGATTACAGGCATAAGCCACGATGCCTGGCCTT
TGTTTTTCACTCTTCTCACTCCCTGAAAGGCATCGTGGGGAGAGGGTGAGTCACTGGACCA
AGTCCTAGAGAACCAGTATCTATTCTTATTCTCAACACATCACCCACGTGACCTGAGC
AAGCCACATACACCCTGGGCCCTAGTTTTATCATCTGTGAAATTAGGGGAAACATAGGT

21787 GGGTGCAATGGTTCACACCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGCGGACCAC
CTGAGGTGAGGATTTGAGACCAGCTGGCCAAACATGGCGAAACCCCGTCTCTACTAAAA
GCACAAAAATTAGCCAGGCGTAGTGGTGCATGCCTGTAGTCCCAGCTACTCGGAAGTCTG
AGGCATGAGAATCACTTGAACCTGGGAGGCAGATGTTGCAGTGAGCCGAGATCGTGCCAC
TGCACTCCAGCTTGGGTGACAGAGCTAGACTGTCTCAAAAAACAAACAAACAAACAAAC
[-, A, C]
TAAAAGATATGTGGATATGAGGGATCACCATCCCCATAGGGCCCTGGATTAACACCACC
CCACCAATGCCCTGAATTAAGAAAGAACAGATGACTAGGTTTGAGAAATCTGGCTTTGG
GTCTATGAGAAGTAGTGTCTCTTTGTGCCTCTTCCATTCTTTTGACATTGAGCTCC
ATGGTGCTCTGAATCCGTCTCTCAGTGTCTGATGGCAGGTGGGACAGATTAGAAAATAG
AGCTGGAGCCACAGAGATTTGGCAGACTGATTCGGTGCCCTCTTGGAAATCTCCAGCACA

22264 CTCCATGGTGCTCTGAATCCGTCTCTCAGTGTCTGATGGCAGGTGGGACAGATTAGAAA

FIGURE 3Q

ATAGAGCTGGAGCCACAGAGATTTGGCAGACTGATTTGCGTGCCCTCTTGAATCTCCAG
CACATTCAAAAAGCCTGGATAGGACCAAAATAGCTTATCAACGTGAGAAAGGACTTCAG
AGCTTGTCTACTGCCAACCTCATTTTACCCAATGAGGAAAGTGAAGCTATTAGGGGGCG
AGGGACACGTGGAAGGTACACAGCACACAGGAGGTGATTACATGTAGATTTACGACC
[T,C]
GCTCCTGCCACGCTGGACTGGTTCACCTCCTAGGCTGACCTGCCTCTCCCCTGTTACA
CACACTCTCGCACACACACACACACACACACAGGTGCTTTGTTCTGGCCAGG
GGTTCCTAGGGTCACCTCTTGGTTGCAGCCACTGTGACCCCAACTGGTCTAACCTCTCTC
TTCCCCTCCCCTTCTTCTGTGGTTCCTGCAGGAGATGTTGGGAGCATGTTTACGCG
AGAGACCCTGACCGCCTGTGAGGGCTCAGCCAGTCTCATGCTGCCTACAGAGTGCCTG

22338 ACAGAGATTTGGCAGACTGATTTGCGTGCCCTCTTGAATCTCCAGCACATTCAAAAAG
CCTGGATAGGACCAAAATAGCTTATCAACGTGAGAAAGGACTTCAGAGCTTGTCTACTGC
CAACCCTCATTTTACCCAATGAGGAAAGTGAAGCTATTAGGGGGCGAGGGACACGTGGAA
GGTCACACAGCACACAGGAGGTGATTACATGTAGATTTACGACACTGCTCCTGCCACG
TGGACTGGTTCACCTCCTAGGCTGACCTGCCTCTCCCCTGTTACACACACTCTCGCAC
[- ,C,A]
CACACACACACACACACACACACAGGTGCTTTGTTCTGGCCAGGGGTTCTAGGGTCA
CCTCTTGGTTGCAGCCACTGTGACCCCAACTGGTCTAACCTCTCTTCCCCTCCCCT
CCTTCTGTGGTTCCTGCAGGAGATGTTGGGAGCATGTTTACGCAGAGACCCTGACCG
CCCTGTGAGGGCTCAGCCAGTCTCATGCTGCCTACAGAGTGCCTGGCACTTGGGACTTC
CATAAAGGATGAGCCTGGGGTCACAGGGGTGTGGGCGGAGGAAAGTGCATCCCCAGA

23363 CAGGGACCATGTGCTCTCCACACCCAGGAGTCTAGGCCTTGGTAACTATGCGCCCCCGT
CCATCATCCCCAAGGCTGCCCAAACCACTGCTGTGAGCAAGCACATCAGACTCTAGC
CTGGACAGTGGCCAGGACCGTCGAGACCACAGAGTACCTCCCCGGGACAGCCCACTA
AGGTTCTGCTCAGCCTCCTGAAACATCACTGCCCTCAGAGGCTGCTCCCTCCCCTGGA
GGCTGGCTAGAAACCCAAAGAGGGGGATGGGTAGCTGGCAGAATCATCTGGCATCCTAG
[T,C]
AATAGATACCACTTATTCTGCACAAAATTTTGGGAATTCCTCTTTCACCCAGAGACTC
AGAGGGGAAGAGGGTGTCTAGTACCAACACAGGGAAAAACGGATGGGACCTGGGCCCAGACA
GTCCCCCTTGACCCAGGGCCCATCAGGAAATGCCTCCCTTGGTAAATCTGCCTTATC
CTTCTTTACCTGGCAAAGAGCCAATCATGTTAACTCTTCTTATCAGCCTGTGGCCAGA
GACACAATGGGGTCTTCTGTAGGCAAAGGTGGAAGTCTCCAGGGATCCGCTACATCCC

23688 AAATTTTGGGAATTCCTCTTTCACCCAGAGACTCAGAGGGGAAGAGGGTGTAGTACC
AACACAGGGAAAAACGGATGGGACCTGGGCCAGACAGTCCCCCTTGACCCAGGGCCCAT
CAGGAAATGCCTCCCTTGGTAAATCTGCCTTATCCTTCTTACCTGGCAAAGAGCCAA
TCATGTTAACTCTTCTTATCAGCCTGTGGCCAGAGACACAATGGGGTCTTCTGTAGG
CAAAGGTGGAAGTCTCCAGGGATCCGCTACATCCCCTAACTGCATGCAGATGTGGAAG
[G,A]
GGCTGATCCAGATTGGGTCTTCTGCACAGGAAGACTCTTTAACACCCTTAGGACCTCAG
GCCATCTTCTCCTATGAAGATGAAAATAGGGGTTAAGTTTTCCATATGTACAAGGAGGTA
TTGAGAGGAACCCCTACTGTTGACTTGAAAATAAATAGGTTCCATGTGTAAGTGTTTTGTA
AAATTTTCACTGGAAATGCACAGAAAATCTTCTGGCCTCTCATCACTGCTTTTCTCAAGCT
TCTTCAGCTTAACAACCCCTTCCCTAACAGGTTGGGCTGGCCCAGCCTAGGAAAACATCC

24210 TCACTGCTTTTCTCAAGCTTCTTCACTTAACAACCCCTTCCCTAACAGGTTGGGCTGGC
CCAGCCTAGGAAAACATCCCCATTTCTAACTTCAGCCAGACCTGCGTTGTGTGTCTGTGT
GTTGAGTGAGCTGGTCAGCTAACAAAGTCTTCTTAGAGTTAAAGGAGGGGGTGTGGCCAA
GAGCCAACACATTCTTGGCCCAGGAGCATTGCTTTTCTGTGAATTCATTATGCCATCTGG
CTGCCAATGGAACCTCAAACTTGGAAGGCGAAGGACAATGTTATCTGGGATTACCGTGC
[A,C]
CAGCACCCAGAGTGCCAAATTCAGGAGGACAAGAGCCTTAGCCAATGACAACTCACTCT
CCCCTACTCCACCTCCTTCCAAGTCCAGCTCAGGCCAGGAGGTGGGAGAAGGTACAGA
GCCTCAGGAATTTCCAAGTCAAGTCCCCTTTGAACCAAGTATCTAGATCCCCTGAGGAC
TTGATGAAGTGATCCTTAACCCCCAAGTAATCATTAAACCCAGACGCTCAGAACTG
AAGGAGATTGTTGACCCAGTGACCTGGAGTTGAGGCTCAGGGAGAGATCTGCCACATGTC

FIGURE 3R